2020

Computation-Driven Analysis of Model Polyextremo- tolerant Fungus Exophiala dermatitidis: Defensive Pigment Metabolic Costs and Human Applications

Wheaton L. Schroeder
University of Nebraska – Lincoln

Steven D. Harris
University of Manitoba, Winnipeg

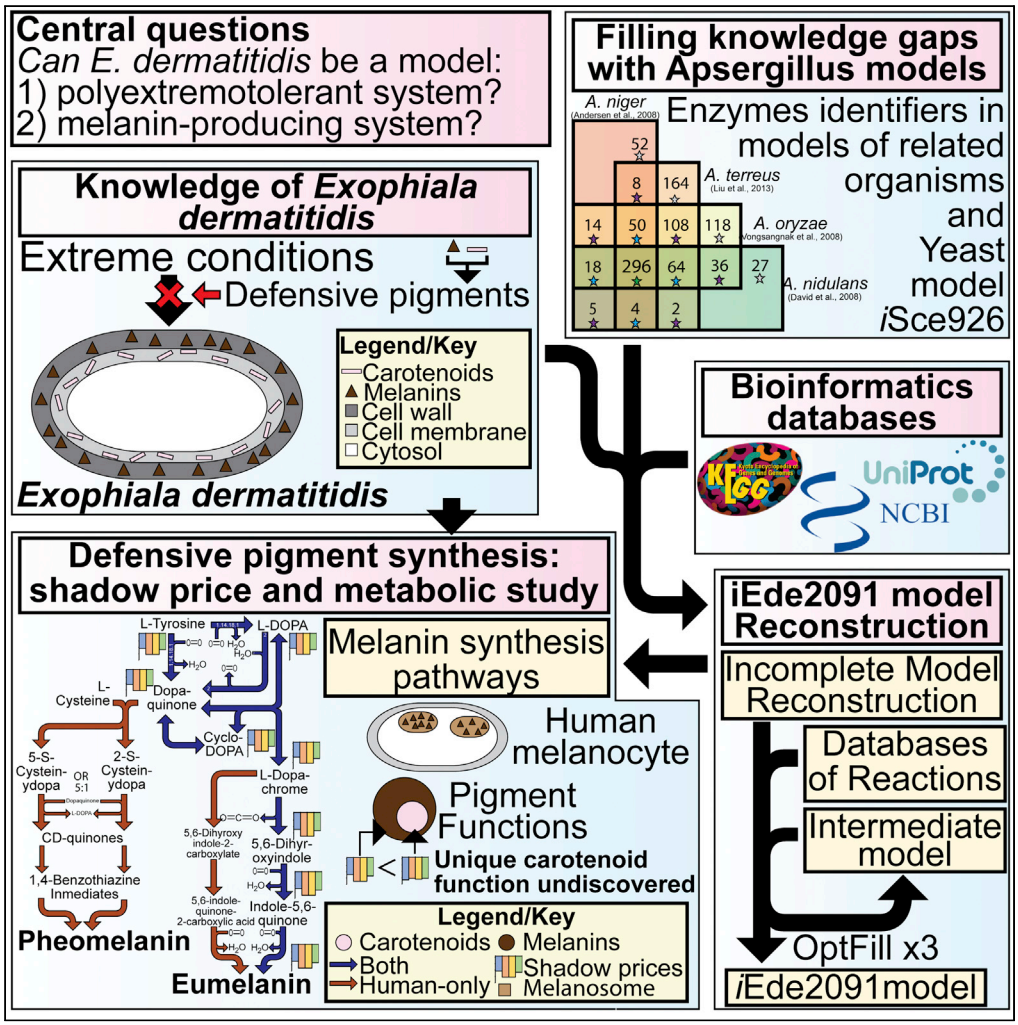
Rajib Saha
National University of Singapore, rsaha2@unl.edu

iScience

CellPress

Article

Computation-Driven Analysis of Model Polyextremotolerant Fungus *Exophiala dermatitidis*: Defensive Pigment Metabolic Costs and Human Applications



Wheaton L. Schroeder, Steven D. Harris, Rajib Saha

rsaha2@unl.edu

HIGHLIGHTS

Exophiala dermatitidis can be a model polyextremotolerant and melaninized organism

Its genome-scale model is reconstructed to study melanin and carotenoid metabolism

The shadow price analysis indicates a potential underexplored role of carotenoids

Comparisons between human and *E. dermatitidis* melanin synthesis are made

Schroeder et al., iScience *23*, 100980
April 24, 2020
https://doi.org/10.1016/j.isci.2020.100980

iScience

CellPress

Article

Computation-Driven Analysis of Model Polyextremotolerant Fungus *Exophiala dermatitidis*: Defensive Pigment Metabolic Costs and Human Applications

Wheaton L. Schroeder,[1] Steven D. Harris,[2] and Rajib Saha[1,3,*]

SUMMARY

The polyextremotolerant black yeast *Exophiala dermatitidis* is a tractable model system for investigation of adaptations that support growth under extreme conditions. Foremost among these adaptations are melanogenesis and carotenogenesis. A particularly important question is their metabolic production cost. However, investigation of this issue has been hindered by a relatively poor systems-level understanding of *E. dermatitidis* metabolism. To address this challenge, a genome-scale model (*i*Ede2091) was developed. Using *i*Ede2091, carotenoids were found to be more expensive to produce than melanins. Given their overlapping protective functions, this suggests that carotenoids have an underexplored yet important role in photo-protection. Furthermore, multiple defensive pigments with overlapping functions might allow *E. dermatitidis* to minimize cost. Because *i*Ede2091 revealed that *E. dermatitidis* synthesizes the same melanins as humans and the active sites of the key tyrosinase enzyme are highly conserved this model may enable a broader understanding of melanin production across kingdoms.

INTRODUCTION

Extremophiles are organisms that can live in extreme conditions of temperature, acidity, alkalinity, or salinity. Studying these organisms not only expands our knowledge on the diversity of life but can also provide significant insights into how organisms adapt to stress, particularly metabolic and regulatory responses. *Exophiala dermatitidis* (hereafter, *Exophiala* or *E. dermatitidis*, also known as *Wangiella dermatitidis*), a highly melanized black fungus and perhaps best known for its *H. sapiens* (hereafter, human) pathogenic properties (Paolo et al., 2006; Poyntner et al., 2016; Sudhadham et al., 2008), is a potential model extremophile system owing to its small genome of 26.4 Mb (Exophiala dermatitidis NIH/UT8656 Genome, 2011) and its demonstrated extremotolerance with respect to temperature (heat and cold) (Paolo et al., 2006; Sudhadham et al., 2008), acidic pH (Sudhadham et al., 2008; Chen et al., 2014), light (Chen et al., 2014; Nosanchuk and Casadevall, 2006; Geis and Szaniszlo, 1984), radiation (Chen et al., 2014; Nosanchuk and Casadevall, 2006; Geis and Szaniszlo, 1984), oxidative stress (Chen et al., 2014; Geis and Szaniszlo, 1984), and likely tolerance to toxic heavy metals (Nosanchuk and Casadevall, 2006), harmful aromatic compounds (Moreno et al., 2018), various toxins (Moreno et al., 2018), antimicrobial compounds (Nosanchuk and Casadevall, 2006), and other stressors (nutrient, osmotic, and mechanical) (Moreno et al., 2018). The ability of *Exophiala* to adapt to most of these conditions seemingly results from two classes of defensive pigments: melanins, a class of pigments consisting of six-carbon ring monomers, and carotenoids, a class of polyisoprenoid pigments. *Exophiala* can produce three different types of melanin: (1) 1,8-dihydroxynaphthalene melanin (hereafter, DHN-melanin), also called naphthalene melanin, (2) DOPA-melanin, also known as eumelanin (Ito and Wakamatsu, 2011), and (3) pyomelanin. Among these, DHN-melanin and pyomelanin are generally produced by fungi (Solano, 2014) including *Exophiala*, whereas eumelanin is produced by both fungi and animals, including humans (Ito and Wakamatsu, 2011; Solano, 2014). The combination of its small genome (Exophiala dermatitidis NIH/UT8656 Genome, 2011), its ability to be cultured as yeast cells (Chen et al., 2014; Ohkusu et al., 1999), and production of eumelanin (Ito and Wakamatsu, 2011) makes *Exophiala* a potential model organism for human melanocytes, the cells in humans that produce melanins. Melanocytes are specialized cells in humans that are found primarily in the skin, which produce pheomelanin and eumelanin in specialized subcellular organelles called melanosomes.

A Genome-Scale Model (GSMs) is a mathematical reconstruction of the metabolism of a target organism, generally accomplished using genomic and biochemical databases to define the sets of gene-protein-reaction (GPR) links (Thiele and Palsson, 2010). GSMs have become an indispensable tool of systems biology

[1]Department of Chemical and Biomolecular Engineering, University of Nebraska – Lincoln, Lincoln, NE 68588, USA

[2]Department of Biological Sciences, University of Manitoba, Winnipeg, MB R3T 2N2, Canada

[3]Lead Contact

*Correspondence: rsaha2@unl.edu

https://doi.org/10.1016/j.isci.2020.100980

in a wide variety of applications (Thiele and Palsson, 2010), with perhaps the most common applications being the overproduction of a native metabolite (Khodayari et al., 2015; Lin et al., 2017; Zhang et al., 2016; Feist and Palsson, 2008) or engineering of metabolism to produce a non-native metabolite (Feist and Palsson, 2008; Gudmundsson et al., 2017). Other uses of GSMs have also been to characterize open reading frames (ORFs), determine gene essentiality, and evolutionary studies in *Escherichia coli* (Feist and Palsson, 2008); investigate the Warburg effect and drug screenings in human cancer cells (Yizhak et al., 2015); study interactions among members of a microbial community (Stolyar et al., 2007; Magnúsdóttir et al., 2016); and investigate plant metabolism under stress (Cheung et al., 2013; Williams et al., 2010; Cramer et al., 2011). Hence, the reconstruction of a GSM of *Exophiala* can be a useful tool to investigate its potential as a model organism both for polyextremotolerant organism and for human melanocytes. However, GSMs are challenging to reconstruct for under-studied organisms such as *Exophiala*, where only approximately 43% of genes have some level of annotation (not including hypothetical or putative proteins) and less than 5% of genes are annotated with Enzyme Commission (EC) numbers, which might be used to establish GPR links (Exophiala dermatitidis NIH/UT8656 Genome, 2011; Exophiala dermatitidis (strain ATCC34100/CBS 525.76/NIH/UT8656) (2018) UniProt2.). This lack of annotations often leaves large gaps in metabolic reconstructions, which requires further scrutiny. One tool that we recently have developed is OptFill (Schroeder and Saha, 2020), which performs whole-model thermodynamically infeasible cycle (TIC) free gapfilling. TICs are detrimental to GSMs as they result in the reporting of unrealistic flux results, cause difficulties in using dual formulations of optimization problems (such as in this work), and can make energy costs such as ATP maintenance meaningless (Schroeder and Saha, 2020). OptFill works by first identifying possible TICs that can occur between a database of functionalities proposed to fix the gaps in the model and the model itself. Then the reaction flux in the direction that would allow a TIC is excluded in the second step of OptFill, which attempts to maximize the number of model reactions fixed by adding new reactions (Schroeder and Saha, 2020). Ultimately, this allows for the maximization of model connectivity while minimizing new functionalities added to the model, as well as opportunity to hypothesize functions for un- or poorly annotated genes through the concurrent use of tools such as BLASTp (Altschul et al., 1997, 2005). Through the process of reconstructing a GSM, metabolic pathways are thoroughly investigated, particularly those related to the subjects of the study, in this case defensive pigments. In addition, this reconstruction provides the basis for comparison between humans and *Exophiala*, which when supplemented with sequence alignment tools such as COBALT (Papadopoulos and Agarwala, 2007) can provide initial comparisons for determining the suitability of *E. dermatitidis* as a model organism.

Once a GSM is reconstructed, optimization-based tools of analysis may be applied to investigate *E. dermatitidis* as a model polyextremotolerant organism. These tools include those that can analyze base metabolism, such as Flux Balance Analysis (FBA) (Orth et al., 2010) and Flux Variability Analysis (FVA) (Gudmundsson and Thiele, 2010); tools that can aid in redesigning metabolism for optimization of a desired phenotype, such as OptKnock (Burgard et al., 2003) and OptForce (Burgard et al., 2003); and tools that elucidate potentially non-intuitive relationships in metabolism such as Flux Coupling Analysis (FCA) (Burgard et al., 2004). This work uses the standard measure of flux of mmol per gDW per h (Orth et al., 2010; Thiele and Palsson, 2010; Maranas and Zomorrodi, 2016). All optimization problems have primal and dual forms, both of which can be enlightening about the problem solution, particularly a quantity determined from the dual problem called the shadow price. The shadow price associated with variable *i* is defined as the reduction in the optimization objective caused by producing one more unit of *i*. Generally, shadow price is used in an economic sense to define the cost of some process in terms of currency; however, this metric can also be applied to the cost of some biological objective (e.g., growth) owing to increasing production of a metabolite, such as a defensive pigment, by one unit. This can be determined using dual formulation of the FBA problem. The cost of producing melanins and carotenoids by *E. dermatitidis* and the associated shadow prices, in particular, have not yet been investigated in this manner.

In this work, a draft GSM of *Exophiala dermatitidis* was first reconstructed from annotated genome of *E. dermatitidis* and an enzyme consensus between four GSMs from a related genus, *Aspergillus*, namely, *A. niger* (Andersen et al., 2008), *A. nidulans* (David et al., 2008), *A. oryzae* (Vongsangnak et al., 2008), and *A. terreus* (Liu et al., 2013). Enzymes used in these *Aspergillus* GSMs (Andersen et al., 2008; David et al., 2008; Vongsangnak et al., 2008; Liu et al., 2013) were used in conjunction with bidirectional BLASTp analyses to hypothesize characterizations of open reading frames. In general, the bidirectional BLASTp analyses assigned EC numbers, and the metabolic functionalities that accompany those numbers, to

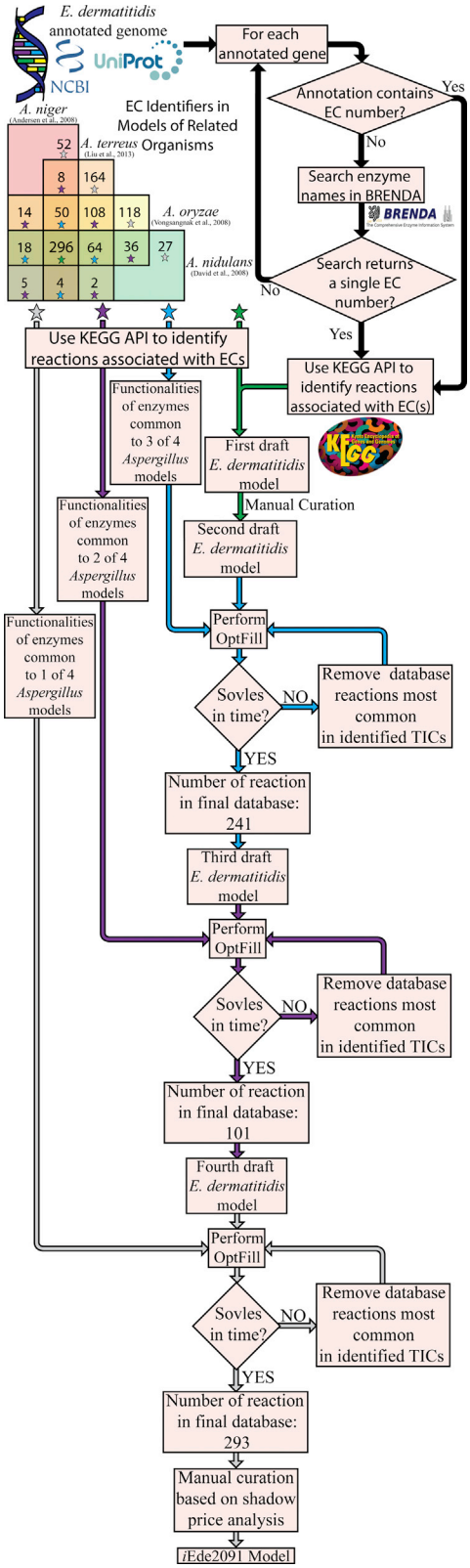

E. dermatitidis
annotated genome
NCBI
UniProt
For each
annotated gene
Annotation contains
EC number?
Yes
No
Search enzyme
names in BRENDA
BRENDA
Search returns
a single EC
number?
No
Yes
EC Identifiers in
Models of Related
Organisms
A. niger
A. terreus
A. oryzae
A. nidulans
52
8
164
14
50
108
118
18
296
64
36
27
5
4
2
Use KEGG API to identify
reactions associated with ECs
Use KEGG API to
identify reactions
associated with EC(s)
KEGG
Functionalities
of enzymes
common
to 3 of 4
Aspergillus
models
First draft
E. dermatitidis
model
Manual Curation
Second draft
E. dermatitidis
model
Functionalities
of enzymes
common
to 2 of 4
Aspergillus
models
Functionalities
of enzymes
common
to 1 of 4
Aspergillus
models
Perform
OptFill
Sovles
in time?
NO
Remove database
reactions most
common
in identified TICs
YES
Number of reaction
in final database:
241
Third draft
E. dermatitidis
model
Perform
OptFill
Sovles
in time?
NO
Remove database
reactions most
common
in identified TICs
YES
Number of reaction
in final database:
101
Fourth draft
E. dermatitidis
model
Perform
OptFill
Sovles
in time?
NO
Remove database
reactions most
common
in identified TICs
YES
Number of reaction
in final database:
293
Manual curation
based on shadow
price analysis
iEde2091 Model

Figure 1. Workflow of *i*Ede2091 GSM Reconstruction

This figure shows the reconstruction workflow of *i*Ede2091, beginning with the annotated genomes from NCBI and UniProt. These gene names taken from these annotated genomes are then used to automatically search the BRENDA database for the associated Enzyme Commission (EC) number. These data were combined with the consensus of enzymes present in the selected *Aspergillus* species GSM reconstructions to form the first draft *E. dermatitidis* GSM model. After manual curation to ensure production of defensive pigments and biomass, this became the second draft *E. dermatitidis* model. Subsequent draft *E. dermatitidis* models were created by using the OptFill tool to fill metabolic gaps using non-consensus *Aspergillus* databases. Once each non-consensus database had been used, the *i*Ede2091 model was complete.

genes already annotated in the NCBI database with non-hypothetical protein names. This draft model next underwent manual and automated curation, the latter through using the tool OptFill (Schroeder and Saha, 2020), to develop the *i*Ede2091 model. *i*Ede2091 was used in computational investigation of the metabolic cost of defensive pigment synthesis through shadow price analysis. This analysis shows that, on both a per-carbon atom and a per-unit (monomer in the case of melanins and molecule in the case of carotenoids), carotenoids are more expensive to produce than melanins. Given that the functions of carotenoids and melanins are generally overlapping, this suggests that carotenoids perform a metabolically valuable protective role that has not been fully explored as of yet, potentially related to absorbance of violet and blue visible light. Finally, the potential of *Exophiala* as a model eumelanin-producing organism, particularly with respect to human eumelanin production in melanocytes, was investigated based on similarity of metabolic pathways and tyrosinase enzyme sequence similarity. This analysis showed that key amino acid residues are conserved in tyrosinase between *Exophiala* and humans, including residues whose mutations are associated with oculocutaneous albinism A1 (OCA1), which suggests *Exophiala* may be used as a model of human eumelanin-production.

RESULTS

Reconstruction of First Draft *E. dermatitidis* Model

In this work, the first draft GSM of *E. dermatitidis* was reconstructed using logical Gene-Protein-Reaction (GPR) links to determine the set of metabolic reactions which occur in an organism using publicly available data such as NCBI and UniProt annotated genomes. This initial reconstruction was necessarily incomplete owing to incomplete genome annotation, in that only approximately 43% of genes were annotated and less than 5% had some Enzyme Commission (EC) number annotation (Exophiala dermatitidis NIH/UT8656 Genome, 2011; Exophiala dermatitidis (strain ATCC34100/CBS 525.76/NIH/UT8656) (2018) UniProt2.). EC numbers were used to establish the GPR links, and therefore automated exploration of BRENDA was used to address this incompleteness and to retrieve more EC numbers; see Figure 1 and Transparent Methods for more details of this procedure. From this, approximately 20% of genes were linked to some EC numbers. These proteins were then localized to their respective subcellular compartment through use of the CELLO subcellular localization tool (Yu and Lin, 2004), the results of which can be found in Table S1. This still left major metabolic gaps; therefore, in addition to genome annotation data, a core set of Enzyme Commission (EC) numbers were identified by being common to GSM models of four strains of a closely related genus (*Aspergillus*), *A. niger* (Andersen et al., 2008), *A. nidulans* (David et al., 2008), *A. oryzae* (Vongsangnak et al., 2008), and *A. terreus* (Liu et al., 2013), hereafter referred to as a the full consensus of *Aspergillus* enzymes. These *Aspergillus* models were chosen as they were the phylogenetically the closest species (Schoch et al., 2009) for which metabolic models were available. This work was limited to using the *Aspergillus* species models in that the next-closest fungi with GSMs published are at the phylum level, for example, *Yarrowia* and *Saccharomyces* species, which are quite phylogenetically distant. Furthermore, all four *Aspergillus* species considered here have larger genome than *E. dermatitidis* allowing for greater genome coverage, whereas model Ascomycetes like *S. cerevisiae* and *Y. lipolytica* have smaller genomes. This restriction resulted in a more conservative metabolic reconstruction than might have otherwise been created in addition to limiting the number of reactions in the database for OptFill applications. This also limited the number of OptFill applications, as each new model considered would require one additional application. The full consensus of *Aspergillus* enzymes included 310 EC numbers in total. ECs already identified in *E. dermatitidis* were removed from the list of EC numbers belonging to the consensus of all four *Aspergillus* models (Andersen et al., 2008; David et al., 2008; Vongsangnak et al., 2008; Liu et al., 2013), leaving 56 unique ECs. These 56 EC numbers were converted to metabolic functionalities and added to the existing draft model of *E. dermatitidis* as a set of functionalities likely common to these closely related melanized fungal species (Schoch et al., 2009). See method and the

GitHub "*E_dermatitidis*_model" repository (https://doi.org/10.5281/zenodo.3608172) for how this was accomplished. Steps taken in reconstruction can be found in greater detail in Table S2.

Bidirectional BLASTp of Full Consensus *Aspergillus* Enzymes onto *E. dermatitidis*

The list of 56 ECs common to *Aspergillus* models but not identified in *Exophiala* were subjected to a bidirectional BLASTp against the *Exophiala* genome. This was accomplished through the Bidirectional BLAST Program (BBP) developed as part of this work, which can be found in the GitHub "E_dermatitidis_model" repository. The BBP program performs forward and backward BLASTp of amino acid sequences, taken from related species, encoding target ECs against a target genome in order to provide evidence for the presence of certain functionalities. The result of the BBP program when applied to the *Aspergillus* consensus ECs (Table S3) was that 39 of 56 (69.6%) consensus ECs were identified in *Exophiala* with 169 unique bidirectional matches using conservative thresholds for the expect (1E-30) and percent positive substitution (60%) values. Many of these matches were between sequences annotated similarly in the reference *Aspergillus* species and *Exophiala*. Examples include annotations in *Aspergillus* species such as "xylulokinase," "2-aminoadipate transaminase," and "phosphoadenylyl-sulfate reductase (thioredoxin)" matching to annotations in *Exophiala* of "D-xylulose kinase A," "aromatic amino acid aminotransferase I," and "phosphoadenosine phosphosulfate reductase," respectively. Other matches assigned EC number to multi-functional enzymes such as the "pentafunctional AROM polypeptide" being assigned to EC numbers 1.1.1.25, 2.5.1.19, and 4.2.1.10 based on strong sequence similarity to specific enzymes such as shikimate dehydrogenase, 3-phosphoshikimate 1-carboxyvinyltransferase, and 3-dehydroquinate dehydratase, respectively. In addition, a total of 22 bidirectional matches to protein sequences currently annotated as "hypothetical proteins" were made. These matches to hypothetical proteins mapped four hypothetical *Exophiala* protein sequences to seven EC numbers. The used reference *Aspergillus* sequences of six of these EC numbers, 1.2.1.38, 2.7.2.8, 6.3.3.1, 6.3.4.13, 6.3.4.14, and 6.4.1.2, only produced significant sequence alignment matched to hypothetical proteins in the *Exophiala* genome, indicating *in silico* identification of potentially unknown metabolic functionalities. Particularly important to this study is the identification EC 6.4.1.2, acetyl-CoA carboxylase, which produces malonyl-CoA. Malonyl-CoA is an essential precursor for the synthesis of hydroxylated naphthalene compounds, which, when polymerized, produce DHN-melanin. See Figure 2 for DHN-melanin synthesis pathway with the reaction catalyzed by EC 6.4.1.2, which highlights the importance of this functionality.

From First Draft *E. dermatitidis* Model to Second Draft *E. dermatitidis* Model

Despite the added functionality of the *Aspergillus* full consensus enzyme set and subsequent potential identification of new functionalities in the *Exophiala* genome, there were a number of "holes" in the metabolic reconstruction. These "holes" included lacking full synthesis pathways for defensive pigments and some biomass components. Therefore, the set of enzymes common to three of these *Aspergillus* models (Andersen et al., 2008; David et al., 2008; Vongsangnak et al., 2008; Liu et al., 2013), the latest model of another ascomycete fungus, *Saccharomyces cerevisiae* (*i*Sce926) (Chowdhury et al., 2015), and literature information on fungal melanin synthesis (Chen et al., 2014; Eisenman and Casadevall, 2012; Paolo et al., 2006; Schmaler-Ripcke et al., 2009; Toledo et al., 2017), were used to manually address some metabolic gaps. Once this manual step was complete, the model could produce all required defensive pigments and biomass components and all TICs were addressed. In addition, the model was further refined to make sure that *Exophiala* can grow on carbon sources such as ethanol (Kumar, 2018), glucose (Poyntner et al., 2016; Chen et al., 2014), and sucrose (Dadachova et al., 2007) and to provide the opportunity to study metabolism, specifically pigment costs, under various different growth conditions. Once these objectives had been met, the resulting model was called the second draft *Exophiala* model. The second draft model had no TICs and consists of 1,591 reactions, of which 711 could carry flux and at best can produce 591 metabolites. For more details on the reconstruction of the first and second draft *Exophiala* models, see the Transparent Methods.

From Second Draft *E. dermatitidis* Model to *i*Ede2091

The remainder of the set of enzymes common to three of four *Aspergillus* models was then converted to their metabolic functionalities (see Transparent Methods), for a total of 344 reactions, and used as a database for the application of OptFill to the second draft model of *Exophiala*. Unfortunately, the large number of reactions in the model and database, as well as the large number of potential TICs between database and model, required several iterations of performing OptFill and removing from the database reactions participating in the most TICs identified in the allotted solution time (1 week), until the database

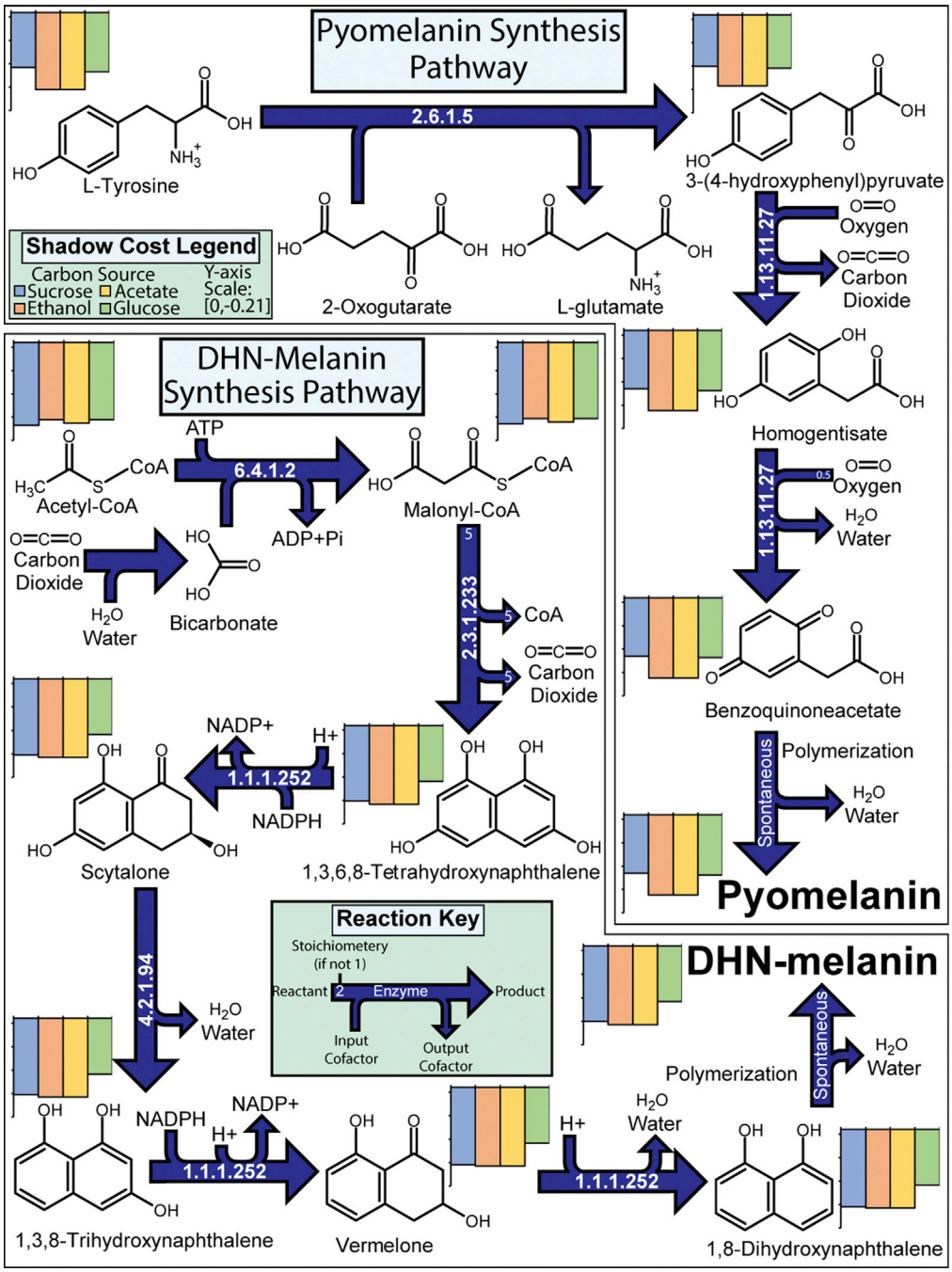


Figure 2. Synthesis Pathways of Pyomelanin and DHN-Melanin in *E. dermatitidis*

This figure shows the synthesis pathways of pyomelanin and DHN-melanin including chemical structures, reaction stoichiometries, catalyzing Enzyme Commission (EC) number, and reaction cofactors.

was reduced to 241 reactions, which allowed reasonable solution times (e.g., under 1 week to produce some gapfilling solutions). This procedure was repeated for the set of enzymes common to two of four *Aspergillus* models and to those unique to one model. This workflow is highlighted in Figure 1. In total, 43 reactions were added to the *Exophiala* model. This resulted in unblocking of a total of 82 reactions and 63 metabolites. Once each solution of this workflow was incorporated, the enzymes linked to filling solutions underwent a bidirectional BLASTp between reference *Aspergillus* sequences and the *Exophiala* genome, to determine the level of genomic support for these added reactions. This procedure was

repeated for the set of enzymes common to two of four *Aspergillus* models and to the set of enzymes belonging to exactly one *Aspergillus* model. The resultant model was designated *i*Ede2091. The *i*Ede2091 model contains 1,661 reactions (of which 824 can carry flux as determined by Flux Variability Analysis), 1,856 metabolites, and 2,091 genes. The set of genes includes those used to build the first draft model (861 genes) and those related to added metabolic functionality from the full consensus of *Aspergillus* model enzymes (33 genes), the set of enzymes common to three of four *Aspergillus* models (21 genes), the set of enzymes common to two of four *Aspergillus* models (2 genes), and the set of enzymes unique to an *Aspergillus* model (18 genes).

Applications of the *i*Ede2091 Model

The *i*Ede2091 model was applied in two investigations. The first is the investigation of the shadow price of defensive pigments to better understand the costs and roles of the defensive pigments in polyextremotolerant systems. The second is the investigation of *Exophiala* melanin synthesis and comparison with that of humans to investigate the feasibility of using *Exophiala* as a model of human melanocytes.

Shadow Price of Defensive Pigments and Their Precursors under Various Growth Conditions

The *i*Ede2091 model was subjected to 36 growth conditions based on the available carbon source (sucrose, ethanol, acetate, or glucose), growth-limiting nutrient (carbon, nitrogen, or sulfur), and rate at which that limiting nutrient was made available to the system (low, moderate, or high). In this study, the growth-limiting nutrient or atom was defined as the nutrient that controls the rate of growth through its scarceness, whereas all other nutrients or atoms are provided in at least three order of magnitude excess. The rate of availability of the growth-limiting nutrient to the organism is also arbitrary because no information appears to be published that would suggest biologically relevant uptake rates for *E. dermatitidis*. Shadow price is the change in the objective value of an optimization problem for one more unit of the desired product. As the model simulations were performed using the objective of maximizing biomass, all shadow prices are negative in value and should be compared against a baseline growth rate of approximately 0.104 h^{-1} for non-stressed *Exophiala* growth in nutrient-limited conditions (Dadachova et al., 2007), since, as can be seen in Table S4, the magnitude of the availability of the limiting resource and the growth rate have no effect on the shadow price. Table S4 shows that, under arbitrarily defined high, medium, and low growth-limiting nutrient availability conditions (corresponding to high, medium, and low growth rates), the shadow price is constant. This was chosen as a baseline for comparison with shadow prices derived from the *i*Ede2091 model because no data are at present available to describe the rate of nutrient uptake by *E. dermatitidis*, which would enable the use of *i*Ede2091 to estimate the growth rate. In the following analyses the per-atom rate of carbon uptake was standardized across the different carbon sources.

Carbon-Limited Conditions

Samples of shadow prices for melanins can be found in Figures 3A and 3B. In general, DHN-melanin is more expensive than eumelanin and pyomelanin both on a per-carbon basis and a per-monomer basis. The higher per-monomer cost of DHN-melanin is due to both the higher per-carbon cost and monomers being composed of ten carbons, as opposed to eight carbons for the other two types of melanin produced by *Exophiala*. As shown in Figures 3A and 3B, not all carbon sources are equally effective in the production of melanins. Generally, melanins are most expensive, in terms of shadow cost, to produce when *Exophiala* is grown using sucrose as a sole carbon source, with the exception of producing eumelanin using acetate as a sole carbon source. For all cases, as suggested by the shadow prices in Figures 3A and 3B, producing one additional mmol·gDW^{-1}·h^{-1} of any melanin monomer would cause *Exophiala* to cease all growth, and even catabolize existing biomass to meet this demand.

In addition to investigating the pigments themselves, an investigation has been made into the shadow cost of precursor molecules to the pigments. Here, a precursor will be defined as molecules consumed by important enzymes related to pigment production or generally agreed upon as the metabolic branching point to pigment synthesis and all molecules "downstream" of that point. For instance, since tyrosinase is considered important in eumelanin synthesis, tyrosine and all molecules in eumelanin synthesis after tyrosine are considered eumelanin precursors. In this work, these pigment precursors have been included in Figures 2, 4, S1, and S2. With respect to the melanin precursors, per-carbon atom cost of the precursors is generally lower than that of the melanins that these produce. Furthermore, precursor per-carbon atom shadow price is generally consistent from the point at which melanin synthesis pathways branch

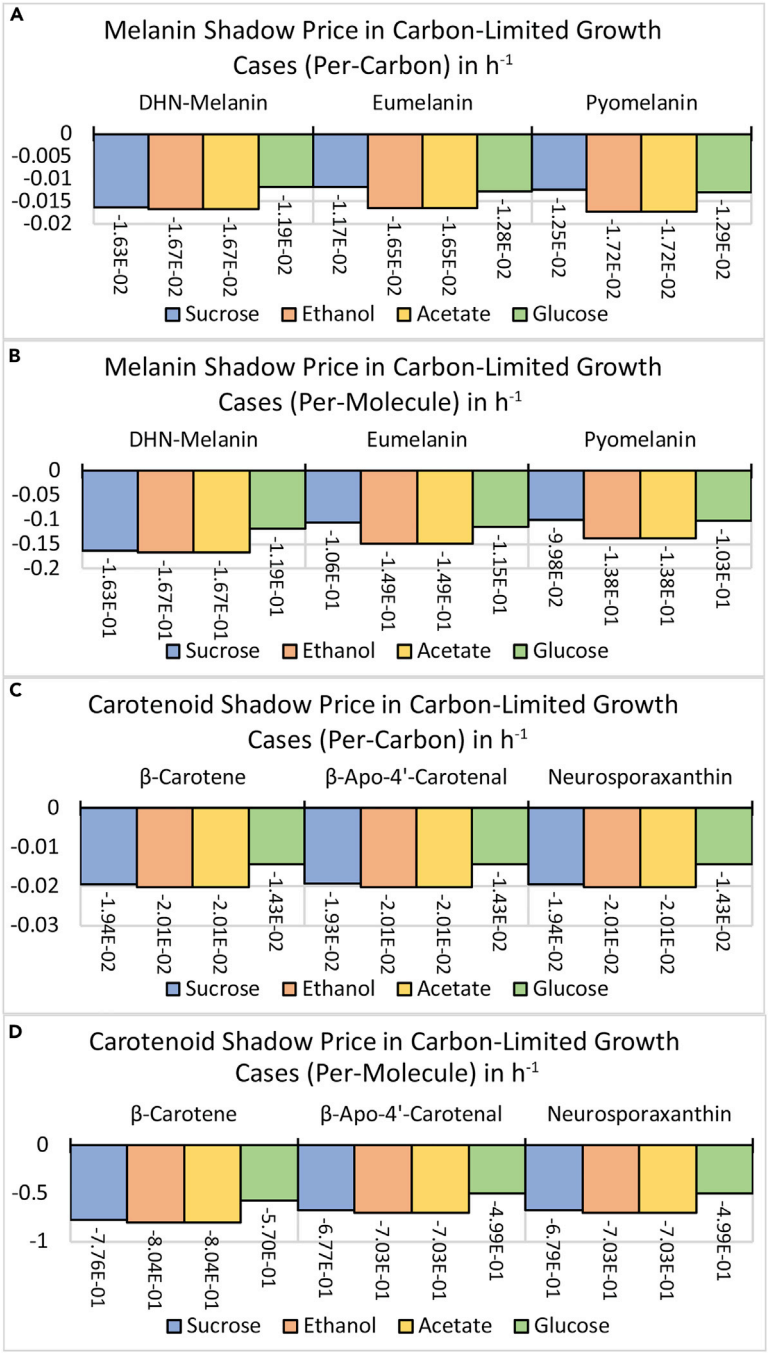


Figure 3. Shadow Prices of *E. dermatitidis* Pigments

This figure shows bar graphs of *E. dermatitidis* defensive pigment shadow prices under carbon-atom limited conditions, using four different carbon sources, on per-limited atom and per-unit basis.

(A) Per-carbon atom shadow costs of the three melanins producible by *E. dermatitidis* under various carbon-limited growth conditions.

(B) Per-monomer shadow costs of the three melanins producible by *E. dermatitidis* under various carbon-limited growth conditions.

(C) Per-carbon atom shadow costs of the three carotenoids that are modeled to constitute *E. dermatitidis* biomass under various carbon-limited growth conditions.

(D) Per-molecule shadow costs of the three carotenoids which are modeled to constitute *E. dermatitidis* biomass under various carbon-limited growth conditions.

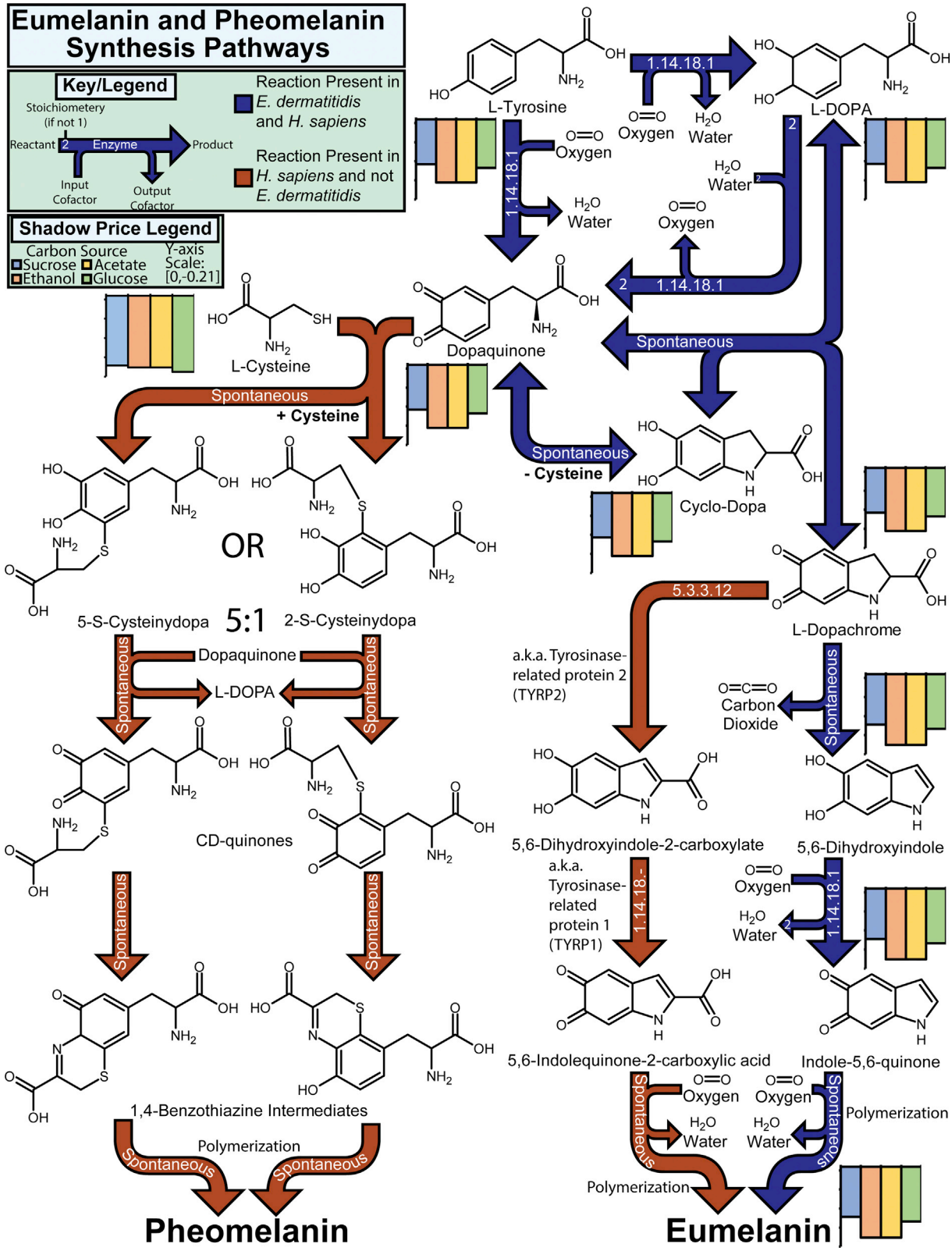

Eumelanin and Pheomelanin Synthesis Pathways
Key/Legend
Stoichiometery (if not 1)
Reactant
Enzyme
Product
Input Cofactor
Output Cofactor
Reaction Present in E. dermatitidis and H. sapiens
Reaction Present in H. sapiens and not E. dermatitidis
Shadow Price Legend
Carbon Source
Sucrose
Acetate
Ethanol
Glucose
Y-axis Scale: [0,-0.21]
L-Tyrosine
1.14.18.1
L-DOPA
Oxygen
Water
L-Cysteine
Dopaquinone
Spontaneous
+ Cysteine
- Cysteine
Cyclo-Dopa
OR
5-S-Cysteinydopa
5:1
2-S-Cysteinydopa
L-Dopachrome
5.3.3.12
a.k.a. Tyrosinase-related protein 2 (TYRP2)
Carbon Dioxide
CD-quinones
5,6-Dihydroxyindole-2-carboxylate
5,6-Dihydroxyindole
a.k.a. Tyrosinase-related protein 1 (TYRP1)
1.14.18.-
5,6-Indolequinone-2-carboxylic acid
Indole-5,6-quinone
1,4-Benzothiazine Intermediates
Polymerization
Pheomelanin
Eumelanin

Figure 4. Synthesis Pathways of Eumelanin and Pheomelanin in Humans and *E. dermatitidis*

This figure shows the synthesis pathways of eumelanin and pheomelanin including chemical structures, reaction stoichiometries, catalyzing Enzyme Commission (EC) number, and reaction cofactors in humans (brown and blue arrows) and *E. dermatitidis* (blue arrows). The major difference between these species' eumelanin synthesis pathways is the presence of tyrosine-related proteins (TYRPs) in humans that catalyze the reactions indicated by brown arrows. In both species, the key initiating enzyme is tyrosinase, Enzyme Commission 1.14.18.1, which catalyzes the initial steps of eumelanin synthesis. A deficiency in tyrosinase activity may result in oculocutaneous albinism A1 in humans. The second type of human melanin, pheomelanin, is largely produced by spontaneous reactions beyond the tyrosinase-catalyzed production of dopaquinone. The branching of eumelanin and pheomelanin production is accomplished by the presence or absence of cysteine where dopaquinone is concentrated. This suggests that pheomelanin may be inducible in *E. dermatitidis*.

from other metabolic pathways (the branchpoint being the starting point of the syntheses depicted in Figures 2 and 4). One example can be clearly seen in the DHN-melanin synthesis pathway with 1,3,6,8-tetrahydroxynaphthalene, scytalone, and 1,3,8-trihydroxynaphthalene all having the same shadow cost. This consistency is not seen in those molecules more proximal to core metabolism such as acetate, ATP, CTP, and requisite amino acids to produce these precursors (such as tyrosine and cysteine). The shadow cost of melanin pigments and their precursors is similar between ethanol and acetate growth conditions. This is because nearly the same set of reactions metabolize both these carbon sources, with the primary difference being the generation of two molecules of NADH in the catalysis of ethanol to acetate. This has no effect on the shadow price of molecules such as tyrosinase but has some effect in the shadow price of carotenoids (ethanol-grown *E. dermatitidis* has a lower shadow price for carotenoids, see Figures S1–S3). It can be noted that, in the shadow prices of melanins and their precursors, these molecules are generally cheaper to produce when grown on sucrose or glucose substrates. This is primarily due to the fact that the precursors to tyrosine synthesis, namely, d-erythrose-4-phosphate (with its own precursors of d-glyceraldehyde-3-phosphate and beta-d-fructose-6-phosphate) and phosphoenolpyruvate, are part of (or proximal to) the glycolysis/gluconeogenesis pathway. From sucrose or glucose, glycolysis is performed to produce these tyrosine precursors. On the other hand, from acetate and ethanol, gluconeogenesis is performed to produce these tyrosine precursors. Gluconeogenesis requires more energy than glycolysis; therefore, the shadow cost of tyrosine-derived pigments is greater for *E. dermatitidis* when grown on acetate or ethanol in comparison with growth on sucrose or glucose.

The per-carbon atom shadow prices of the three carotenoids that are a part of *Exophiala* biomass as modeled in *i*Ede2091, namely, beta carotene, β-apo-4′-carotenal, and neurosporaxanthin, are approximately equivalent; see Figures 3C and 3D. Synthesis pathways used by *E. dermatitidis* to produce carotenoids, as well as the shadow prices of carotenoid precursors, can be found in Figures S1–S3. As the per-carbon atom shadow costs are approximately equivalent, the per-molecule differences in shadow cost are due to the difference in number of carbon atoms in the carotenoid molecules, as beta carotene contains 40 carbon atoms, whereas the other two carotenoid compounds contain 35 carbon atoms. Essentially, carotenoids are more expensive for the cell to produce than are melanins on a per-carbon atom basis.

Nitrogen- and Sulfur-Limited Conditions

The nitrogen source used by the model is ammonia and, as with the carbon-limited conditions, the availability of the growth-limiting nutrient has no effect on shadow cost. In this analysis, metabolites that do not contain nitrogen, including DHN melanin, pyomelanin, and all three investigated carotenoids, have no shadow cost under nitrogen-limited conditions. This makes sense in that all other atoms are provided to the system in excess; therefore, utilizing more of those excess atoms would not hamper biomass production. As such, the only melanin compound that has a shadow cost in these conditions is eumelanin, whose monomers contain a single nitrogen. In nitrogen-limited conditions, the per-nitrogen atom shadow cost is approximately 41 times higher than that of the per-carbon atom cost. The reasons for this are likely 2-fold. First, far less nitrogen is needed by *Exophiala* to produce biomass than carbon (approximately 9.1:1 C:N in the biomass pseudomolecule). Second, not all nitrogen uptaken can be used by *Exophiala* and utilization of nitrogen is less efficient than utilization of carbon. For instance, waste nitrogen is excreted in a nitrogen compound containing four nitrogen atoms (in urate), as opposed to the majority of waste carbon being expelled as carbon dioxide.

Similarly, in the cases where sulfur is the nutrient limiting model growth, compounds that contain no sulfur atoms have no shadow cost, including all defensive pigments studied. Therefore, only melanin precursors have a shadow cost under these conditions, which includes coenzyme A (CoA), its precursors, and all molecules containing CoA such as malonyl-CoA and acetyl-CoA. These compounds have relatively high

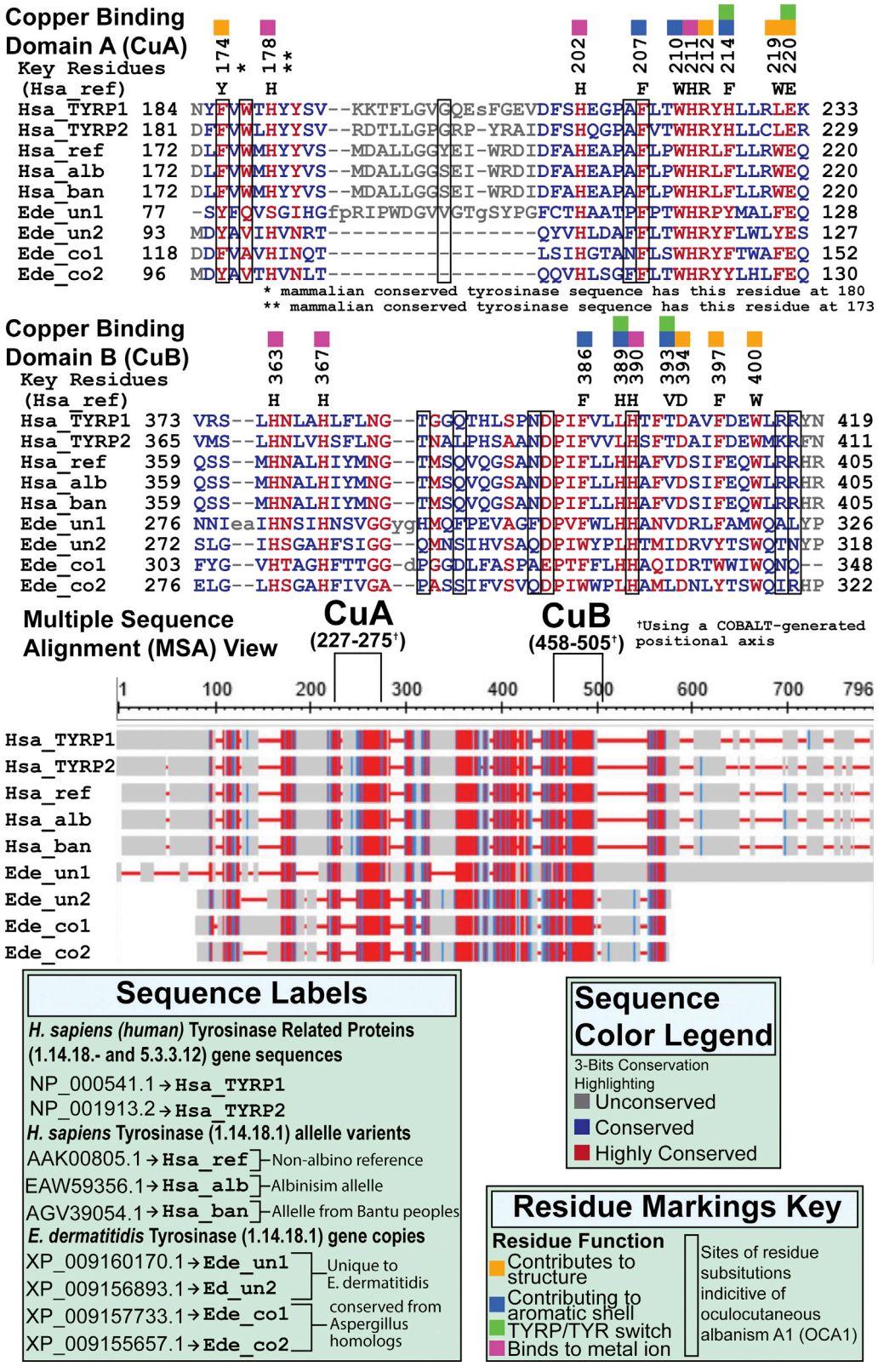

Copper Binding Domain A (CuA)
Key Residues (Hsa_ref)
Y H H F WHR F WE
Hsa_TYRP1 184 NYFVWTHYYSV--KKTFLGVGQEsFGEVDFSHEGPAFLTWHRYHLLRLEK 233
Hsa_TYRP2 181 DFFVWLHYYSV--RDTLLGPGRP-YRAIDFSHQGPAFVTWHRYHLLCLER 229
Hsa_ref 172 DLFVWMHYYVS--MDALLGGYEI-WRDIDFAHEAPAFLPWHRLFLLRWEQ 220
Hsa_alb 172 DLFVWMHYYVS--MDALLGGSEI-WRDIDFAHEAPAFLPWHRLFLLRWEQ 220
Hsa_ban 172 DLFVWMHYYVS--MDALLGGSEI-WRDIDFAHEAPAFLPWHRLFLLRWEQ 220
Ede_un1 77 -SYFQVSGIHGfpRIPWDGVVGTgSYPGFCTHAATPFPTWHRPYMALFEQ 128
Ede_un2 93 MDYAVIHVNRT----------QYVHLDAFFLTWHRYFLWLYES 127
Ede_co1 118 DDFVAVHINQT----------LSIHGTANFLSWHRYFTWAFEQ 152
Ede_co2 96 MDYAVTHVNLT----------QQVHLSGFFLTWHRYYLHLFEQ 130
* mammalian conserved tyrosinase sequence has this residue at 180
** mammalian conserved tyrosinase sequence has this residue at 173
Copper Binding Domain B (CuB)
Key Residues (Hsa_ref)
H H F HH VD F W
Hsa_TYRP1 373 VRS--LHNLAHLFLNG--TGGQTHLSPNDPIFVLLHTFTDAVFDEWLRRYN 419
Hsa_TYRP2 365 VMS--LHNLVHSFLNG--TNALPHSAANDPIFVVLHSFTDAIFDEWMKRFN 411
Hsa_ref 359 QSS--MHNALHIYMNG--TMSQVQGSANDPIFLLHHAFVDSIFEQWLRRHR 405
Hsa_alb 359 QSS--MHNALHIYMNG--TMSQVQGSANDPIFLLHHAFVDSIFEQWLRRHR 405
Hsa_ban 359 QSS--MHNALHIYMNG--TMSQVQGSANDPIFLLHHAFVDSIFEQWLRRHR 405
Ede_un1 276 NNIeaIHNSIHNSVGGygHMQFPEVAGFDPVFWLHHANVDRLFAMWQALYP 326
Ede_un2 272 SLG--IHSGAHFSIGG--QMNSIHVSAQDPIWYPLHTMIDRVYTSWQTNYP 318
Ede_co1 303 FYG--VHTAGHFTTGG-dPGGDLFASPAEPTFFLHHAQIDRTWWIWQNQ-- 348
Ede_co2 276 ELG--LHSGAHFIVGA--PASSIFVSVQDPIWWPLHAMLDNLYTSWQIRHP 322
Multiple Sequence Alignment (MSA) View
CuA (227-275†)
CuB (458-505†)
†Using a COBALT-generated positional axis
1 100 200 300 400 500 600 700 796
Hsa_TYRP1
Hsa_TYRP2
Hsa_ref
Hsa_alb
Hsa_ban
Ede_un1
Ede_un2
Ede_co1
Ede_co2
Sequence Labels
H. sapiens (human) Tyrosinase Related Proteins (1.14.18.- and 5.3.3.12) gene sequences
NP_000541.1 → Hsa_TYRP1
NP_001913.2 → Hsa_TYRP2
H. sapiens Tyrosinase (1.14.18.1) allele varients
AAK00805.1 → Hsa_ref — Non-albino reference
EAW59356.1 → Hsa_alb — Albinism allele
AGV39054.1 → Hsa_ban — Allele from Bantu peoples
E. dermatitidis Tyrosinase (1.14.18.1) gene copies
XP_009160170.1 → Ede_un1
XP_009156893.1 → Ed_un2
Unique to E. dermatitidis
XP_009157733.1 → Ede_co1
XP_009155657.1 → Ede_co2
conserved from Aspergillus homologs
Sequence Color Legend
3-Bits Conservation Highlighting
Unconserved
Conserved
Highly Conserved
Residue Markings Key
Residue Function
Contributes to structure
Contributing to aromatic shell
TYRP/TYR switch
Binds to metal ion
Sites of residue subsitutions indicitive of oculocutaneous albanism A1 (OCA1)

Figure 5. Tyrosinase and Tyrosinase-Related Protein Sequence Alignments Between *Humans* and *E. dermatitidis*

This figure shows portions of the sequence alignments performed by NCBI's COBALT tool using the amino acid sequences of human tyrosinase-related protein 1 (Has_TYRP1, accession NP_000541.1), 2 (Has_TYRP2, accession NP_01913.2), a reference allele human tyrosinase sequence (Has_ref, accession AAK00805.1), an oculocutaneous albinism A1 allele (Has_alb, accession EAW59356.1), an allele from an individual of the Bantu peoples (Has_ban, accession AGV39054.1), and reference sequences for the four tyrosinase gene copies of *E. dermatitidis* (Ede_un1, accession XP_009160170.1; Ede_un2, accession XP_009156893.1; Ede_co1, accession XP_009157733.1; and Ede_co2, accession XP_009155657.1). The portions of the alignments shown concern the two parts of the active site of tyrosinase, Copper-Binding Domains A and B (CuA and CuB, respectively). It is shown that all amino acid residues thought to be critical to active site function (key residues) are highly conserved between the two species (García-Borrón and Solano, 2002). Furthermore, many sites where amino acid substitutions are associated with oculocutaneous albinism A1 (residues boxed in black) are conserved between species. Also shown is the Multiple Sequence Alignment (MSA) view, which shows that the active sites and many sequences between the active sites are preserved between the aligned sequences. This further shows that the large differences in sequence lengths between genes are largely due to sequences flanking the active sites.

per-sulfur atom shadow costs, of $-28.73\ h^{-1}$, since each mole of the biomass pseudomolecule contains approximately 0.035 sulfur atoms, indicating that the sulfur needs of *Exophiala* are very low. Therefore, to produce one extra $mmol \cdot gDW^{-1} \cdot h^{-1}$ of a sulfur-containing compound, a large amount of biomass would need to be catabolized.

Comparison of Human and *E. dermatitidis* Melanin Synthesis

The melanin synthesis pathway of *Exophiala* and humans was compared in two ways: first, by the series of reactions that produce human melanins (namely, pheomelanin and eumelanin, see Figure 4), and second, by comparison of the tyrosinase enzymes (see Figure 5).

In building the *i*Ede2091 model, we recognized that fungal melanins are typically transported in exocytic vesicles to the cell surface, where they are then attached to the cell wall (Camacho et al., 2019; Upadhyay et al., 2016). This pathway shares features with that observed in melanocytes, whereby synthesis occurs in specialized melanosomes. Moreover, the *Exophiala* and human pathways to produce the indole-5,6-quinone monomer of eumelanin are identical. Furthermore, the production of pheomelanin in humans appears replicable in *Exophiala* should cysteine be added to the extracellular environment. The 5,6-indolequinone-2-carboxilic acid eumelanin monomer is not producible by *Exophiala* owing to its lack of a tyrosinase-related protein. In investigating the potential for *Exophiala* to produce pheomelanin, the shadow price of cysteine was also investigated. With respect to carbon-limited conditions (see Figure 4), cysteine is more expensive than most other precursors on a per-carbon basis, particularly in cases of growth on sucrose and glucose. With respect to nitrogen-limited growth cases (see Table S4), cysteine is very similar in cost to other amino acids. With respect to sulfur-limited growth cases (see Table S4), the per-sulfur atom cost is high (around $28\ h^{-1}$) and is similar in cost to coenzyme A.

The four tyrosinase gene copies in *E. dermatitidis* where identified through genome annotation. Furthermore, these sequences were used as a BLASTp query against the *E. dermatitidis* genome to confirm that these four were the only tyrosinase gene copies in *E. dermatitidis*. A non-redundant BLASTp analysis was performed by using the tyrosinase amino acid sequences of *Exophiala* as the search sequence against the human genome to determine the sequence similarity. This produced no matches of acceptable expect value (e.g., less than 1E-10), indicating large sequence dissimilarities. However, a COBALT alignment (Papadopoulos and Agarwala, 2007) of the amino acid sequences of three human tyrosinase alleles, human tyrosinase-related proteins, and the four gene copies of *Exophiala* produces more nuanced results. Tyrosinase-related proteins (TYRPs) have the same evolutionary origin as tyrosinase and are still very similar and were therefore included in this analysis (Furumura et al., 1998). The major catalytic difference between TYRPs and tyrosinases is that they act upon L-Dopachrome differently, one producing 5,6-indolequinone-2-carboxylic acid eumelanin monomers and the other producing indole-5,6-quinone eumelanin monomers. Portions of this alignment, namely, the sequences related to the copper-binding domains A (CuA) and B (CuB) that constitute the active side of tyrosinase, are shown in Figure 5 using the 3-bit highlighting method. This method highlights in red aligned residues that have the same or very similar chemical structure, in blue somewhat conserved regions, and in gray unconserved regions. When highlighting key structural (orange squares located above the residue number), functional (blue squares), and active site (purple squares) residues, it appears that these key residues are highly conserved between human tyrosinase-related proteins and tyrosinase and *Exophiala* tyrosinases. Poor BLASTp alignment scores appear to be due to substitutions, deletions, or lack of sequence

conservation of non-critical residues, gaps in less critical regions of tyrosinase (such as residues that are not a part of secondary structures, such as the gap in CuA), and significant differences in enzyme length. This is shown in the Multiple Sequence Alignment (MSA) view shown in Figure 5. As an example of the length differences, human tyrosinase has a primary structure of 529 amino acids, and tyrosinase-related proteins 1 and 2 have structures of 537 and 519 amino acids, respectively, whereas *Exophiala* tyrosinase lengths range from 381 to 614 amino acids. Much of the differences in length are in those sequences upstream of CuA and downstream of CuB; see the sequence identity summary shown in Figure 5. Interestingly, Figure 5 highlights residue mutations that trigger the switch between tyrosinase and tyrosinase-related proteins (green squares at 214, 219, 389, and 393 [García-Borrón and Solano, 2002]). In some gene copies of *Exophiala* tyrosinase, particularly the copy labeled as "Ede_un1," key residues that when mutated cause the switch between tyrosinase and tyrosinase-related protein are particularly well conserved, suggesting that *Exophiala* could be engineered to have a tyrosinase-related protein. Should this additional monomer synthesis pathway be engineered in *Exophiala*, through gene insertion or selective mutation, the melanin synthesis pathways between *Exophiala* and humans could be very similar. In addition to this analysis, a sequence alignment analysis to the Hidden Markov Model (HMM) using the Pfam tool (El-Gebali et al., 2019) was performed. This tool acknowledged the strong sequence similarity of *E. dermatitidis* tyrosinase enzymes with that of the general pattern of tyrosinase enzymes. The results of this analysis can be found in Data S1.

In considering the uses of *Exophiala* as a model system of human melanin production, some amino acid residue positions where residue substitutions are associated with oculocutaneous albinism A1 (OCA1), which accounts for approximately 50% of cases of albinism worldwide and is caused by a non-functional tyrosinase in humans (Kamaraj and Purohit, 2014), are shown in black rectangles (Spritz, 1994) to highlight the potential for *Exophiala* as a model system to study OCA1.

DISCUSSION

In this work, a stoichiometric GSM of *E. dermatitidis* (*i*Ede2091) consisting of 1,661 reactions, 1,856 metabolites, and 2,091 genes was developed in order to investigate *Exophiala* as a potential model organism for extremotolerant fungi and human melanocytes. Several issues were encountered in the metabolic reconstruction. First, the low levels of genome annotation (43% annotated but less than 5% with associated enzyme commission numbers) represented knowledge gaps in the understanding of *Exophiala* metabolism that lead to many gaps and blocked reactions throughout the stages of reconstruction. This was dealt with by using four metabolic models from the related *Aspergillus* genus (Andersen et al., 2008; David et al., 2008; Vongsangnak et al., 2008; Liu et al., 2013) in addition to the OptFill tool (Schroeder and Saha, 2020) for TIC-free gapfilling of models; see Figure 1. The low levels of genome annotation also hindered the ability to create GPR links, which was addressed by using *Aspergillus* protein sequences as enzyme reference sequences for use in BLASTp analyses. This resulted in a large number of previously annotated genes being linked with enzyme commission numbers and the functional identification of four sequences that may not yet have been identified.

In the shadow price investigation of melanins (Figure 3), it was noted that DHN-melanin has a higher per-unit cost than other melanins. This appears to be due simply to the larger number of carbon molecules present in each monomer unit when compared with other melanins (see Figures 2 and 4). Furthermore, the difference between DHN-melanin, eumelanin, and pyomelanin in media where sucrose is the limited carbon source is that the latter two are synthesized from l-tyrosine, whereas DHN-melanin is synthesized from malonyl-CoA. The higher shadow price appears to be due to the higher per-carbon atom cost to produce acetate from sucrose, which is perpetuated through the DHN-synthesis pathway. As shown in Figure 3, both melanin and carotenoid pigments are "cheapest" to produce in carbon-limited cases when glucose is the carbon source. This is due to the lack of preprocessing needed (e.g., other carbon sources may require gluconeogenesis or other metabolic transformations before being shunted to major energy-harvesting pathways).

The changing shadow prices for these defensive pigments under different growth conditions suggest that the profile of pigments (i.e., the type and quantity of defensive pigments) as produced by *Exophiala* varies by nutrient availability. In other words, the "cheaper" defensive pigments may be produced more than the expensive pigments. Having a range of defensive pigments (e.g., three melanin types and various carotenoids) with differing synthesis pathways makes them to be more or less expensive depending on available nutrients. This, in turn, may help minimize the cost of the extremotolerant nature of *Exophiala*

by allowing the organism to preferentially produce the least expensive defensive pigment(s). The relatively high fractions of biomass accounted for by defensive pigments, 1.3 wt% for melanin (Geis, 1981) and 3.5 wt % for carotenoids (Strobel et al., 2009), as well as their high shadow prices suggest that these pigments are continually produced and stockpiled because increasing production of these pigments to meet cell need if the environment were to quickly become extreme is untenable.

The higher per-carbon shadow prices of carotenoids compared with melanins might help to expand the current understanding of the role of carotenoids. First, carotenoids are a secondary line of defense against external extreme conditions as they are deposited in the cell membrane (Chen et al., 2014; Kumar, 2018, whereas melanins are deposited in the cell wall (Chen et al., 2014; Geis, 1981; Szaniszlo, 2002). Second, melanins are known to provide protection against antifungal and antimicrobial compounds (Nosanchuk and Casadevall, 2006; Paolo et al., 2006; Toledo et al., 2017); lytic enzymes (Paolo et al., 2006); heat and cold stress (Paolo et al., 2006; Toledo et al., 2017); rapid freezing (Paolo et al., 2006); ionizing radiation (Kumar, 2018; Dadachova et al., 2007); oxidative stress (Toledo et al., 2017); UV radiation (Toledo et al., 2017); heavy metals (Kumar, 2018; Singh et al., 2013); light (Chen et al., 2014); and immune responses (Chen et al., 2014). Furthermore, several genes related to both eumelanin and DHN-melanin synthesis are upregulated under low pH stress, suggesting that melanins are also produced under pH stress (Chen et al., 2014). At present, it is known that carotenoids protect against stress conditions such as oxidative stress (such as free radicals) (Kumar, 2018; Avalos and Carmen Limón, 2015; Strobel et al., 2009), UV radiation (Avalos and Carmen Limón, 2015; Geis and Szaniszlo, 1984; Kumar, 2018; Strobel et al., 2009), and light (Kumar, 2018; Avalos and Carmen Limón, 2015; Strobel et al., 2009). Each function of carotenoids is already accounted for by melanins. It has been suggested that carotenoids do not play a physiological role in fungi, but rather function as precursors to the synthesis of other biomolecules (Avalos and Carmen Limón, 2015). However, this appears inconsistent with their higher shadow cost in comparison with melanin compounds, which can accomplish the same functions with deposition in the cell membrane and high weight fraction in some fungal species (Strobel et al., 2009). Although several previous works hinted about the possibility of carotenoids having unexplored functions in fungi (Chen et al., 2014; Avalos and Carmen Limón, 2015), this is the first study that provides a computational and systems biology perspective. One study has postulated that perhaps carotenoids protect against light that passes through the melanin in the cell wall (Chen et al., 2014). This seems a likely function as melanin absorbance of electromagnetic radiation is high in the UV spectrum to approximately 400 nm in wavelength and exponentially declines in the wavelength range of 400–500 nm (Mahmood et al., 2015; Ou-Yang et al., 2004), whereas this latter range constitutes the peak absorbance of carotenoids (Yamamoto and Bangham, 1978; Zaghdoudi et al., 2017). Thus, the combination of these two pigments would protect *Exophiala* cell from the UV spectrum through higher-energy visible light (namely, violet and blue light). The high cost of producing carotenoids along with high fraction of cell weight does suggest that the violet and/or blue light is particularly disruptive to some high-value metabolic process in *Exophiala*, which should be further investigated.

In exploring the suitability of *Exophiala* as a model organism for human melanocytes, the sequence alignment results of *Exophiala* and human tyrosinase enzymes show that CuB is the best-conserved portion of tyrosinase active site, through all key amino acids, and therefore likely the essential structures of CuA is also preserved. As tyrosinase is the key enzyme in eumelanin synthesis in both *E. dermatitidis* and human, several residues associated with OCA1 are persevered between the species. Since *Exophiala* has a significantly smaller genome (26.4 Mb compared with 3253.8 Mb for human), *Exophiala* may be used as a model system for human eumelanin production. As OCA1 is the most prevalent type of albinism worldwide, *Exophiala* may be used as a model system for studying causal mechanisms of OCA1 and potentially to identify treatment options. Unfortunately, African populations, where albinism is a more pressing social and health problem (Brilliant, 2015), would benefit less from *Exophiala* eumelanin studies, than Caucasian and Asian populations, as approximately 77% of albinism cases in African populations result from oculocutaneous albinism A2 (OCA2), with most of the remainder being attributed to OCA1. OCA2 is a result of the lack of a tyrosinase transporter proteins, called P protein, which is necessary to transport tyrosinase into human melanosomes (a subcellular compartment dedicated to melanin synthesis in melanocytes) and/or stabilize tyrosinase (Kamaraj and Purohit, 2014), which was not identifiable through *in silico* methods in *Exophiala*, such as through BLASTp or annotated genomes. Therefore, further study of *Exophiala* is warranted to3 identify this transporter protein and improve the potential for *Exophiala* as a model system.

Furthermore, it was determined that one type of melanin produced by human is not produced by *Exophiala*. Pheomelanin is a red-brown to yellow pigment (Ito and Wakamatsu, 2011), and it is likely that *Exophiala* could produce this type of melanin. No additional enzymes are needed to produce pheomelanin beyond that which *Exophiala* already possesses (see Figure 4), rather free cysteine in the location of eumelanin synthesis is required (Ito, 2003). Growing *Exophiala* in a cysteine-rich culture or engineering a cysteine pump to the extracellular space of *Exophiala* could result in pheomelanin production, allowing production of both human melanin types. Alternatively, if *Exophiala* were to provide the cysteine for pheomelanin synthesis, given its high shadow price and the shadow price of dopaquinone, it is reasonable to hypothesize that the resultant pheomelanin would be the costliest melanin produced by *E. dermatitidis*. This is perhaps why *E. dermatitidis* does not natively produce pheomelanin.

Overall, the results of this work suggest several potential interesting *in vivo* follow-up studies that will increase our understanding of extremotolerant fungi using *Exophiala* as a model system. Key predictions arising from the *i*Ede2091 model that are currently being tested include assessing the effects of different carbon sources on melanin and carotenoid accumulation, as well as determining the effects of mutations that abrogate specific metabolic pathways on pigment production. In addition, phenotypic profiling of mutants defective in melanin and/or carotenoid synthesis is underway to better evaluate the roles of each pigment in stress tolerance. Although detailed *in vivo* investigation may be needed to further establish *Exophiala* as a potential model organism for human melanocytes including demonstrating the production of pheomelanin, this work attempts to enable a broader understanding of melanin production across kingdoms.

Limitations of the Study

The major limitation of this study is that no *in vivo* experiments accompany this work; therefore, this work focuses on an *in silico* systems level analysis of metabolism. Furthermore, a number of reactions in the model, approximately 50%, cannot carry flux. This limitation is due in large part to lack of system knowledge that this work was not entirely able to address.

METHODS

All methods can be found in the accompanying Transparent Methods supplemental file.

DATA AND CODE AVAILABILITY

The published article does not include all datasets and code generated or analyzed during this study. The datasets and code generated during this study are available at GitHub in the ssbio/OptFill repository (https://doi.org/10.5281/zenodo.3608172).

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.isci.2020.100980.

ACKNOWLEDGMENTS

This work has been completed utilizing the Holland Computing Center of the University of Nebraska, which receives support from the Nebraska Research Initiative (United States of America). The authors gratefully acknowledge funding from University of Nebraska-Lincoln (United States of America) Faculty Startup Grant to R.S.. S.D.H. wishes to acknowledge support from the Natural Sciences and Engineering Research Council of Canada (NSERC) of Canada Discovery Grants program (Canada).

AUTHOR CONTRIBUTIONS

Conceptualization, W.L.S., S.D.H., and R.S.; Data curation, W.L.S.; Formal analysis, W.L.S.; Funding Acquisition, R.S.; Investigation, W.L.S.; Methodology, W.L.S.; Project administration, R.S.; Resources, R.S. and S.D.H.; Software, W.L.S.; Supervision, R.S.; Validation, W.L.S.; Visualization, W.L.S.; Writing – original draft, W.L.S. and R.S.; Writing – reviewing & editing – W.L.S., S.D.H., and R.S.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: January 21, 2020
Revised: February 28, 2020
Accepted: March 9, 2020
Published: April 24, 2020

REFERENCES

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. *25*, 3389–3402.

Altschul, S.F., Wootton, J.C., Gertz, E.M., Agarwala, R., Morgulis, A., Schäffer, A.A., and Yu, Y.K. (2005). Protein database searches using compositionally adjusted substitution matrices. FEBS J. *272*, 5101–5109.

Andersen, M.R., Nielsen, M.L., and Nielsen, J. (2008). Metabolic model integration of the bibliome, genome, metabolome and reactome of *Aspergillus niger*. Mol. Syst. Biol. *4*, 178.

Avalos, J., and Carmen Limón, M. (2015). Biological roles of fungal carotenoids. Curr. Genet. *61*, 309–324.

Brilliant, M.H. (2015). Albinism in Africa: a medical and social emergency. Int. Health *7*, 223–225.

Burgard, A.P., Nikolaev, E.V., Schilling, C.H., and Maranas, C.D. (2004). Flux coupling analysis of genome-scale metabolic network reconstructions. Genome Res. *14*, 301–312.

Burgard, A.P., Pharkya, P., and Maranas, C.D. (2003). OptKnock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. Biotechnol. Bioeng. *84*, 647–657.

Camacho, E., Vij, R., Chrissian, C., Prados-Rosales, R., Gil, D., O'Meally, R.N., Cordero, R.J.B., Cole, R.N., McCaffery, J.M., Stark, R.E., and Casadevall, A. (2019). The structural unit of melanin in the cell wall of the fungal pathogen *Cryptococcus neoformans*. J. Biol. Chem. *294*, 10471–10489.

Chen, Z., Martinez, D.A., Gujja, S., Sykes, S.M., Zeng, Q., Szaniszlo, P.J., Wang, Z., and Cuomo, C.A. (2014). Comparative genomic and transcriptomic analysis of *Wangiella dermatitidis*, a major cause of phaeohyphomycosis and a model black yeast human pathogen. G3 (Bethesda, Md) *4*, 561–578.

Cheung, C.Y.M., Williams, T.C., Poolman, M.G., Fell, D.A., Ratcliffe, R.G., and Sweetlove, L.J. (2013). A method for accounting for maintenance costs in flux balance analysis improves the prediction of plant cell metabolic phenotypes under stress conditions. Plant J. *75*, 1050–1061.

Chowdhury, R., Chowdhury, A., and Maranas, C.D. (2015). Using gene essentiality and synthetic lethality information to correct yeast and CHO cell genome-scale models. Metabolites *5*, 536–570.

Cramer, G.R., Urano, K., Delrot, S., Pezzotti, M., and Shinozaki, K. (2011). Effects of abiotic stress on plants: a systems biology perspective. BMC Plant Biol. *11*, 163.

Dadachova, E., Bryan, R.A., Huang, X., Moadel, T., Schweitzer, A.D., Aisen, P., Nosanchuk, J.D., and Casadevall, A. (2007). Ionizing radiation changes the electronic properties of melanin and enhances the growth of melanized fungi. PLoS One *2*, e457.

David, H., Ozçelik, I.S., Hofmann, G., and Nielsen, J. (2008). Analysis of *Aspergillus nidulans* metabolism at the genome-scale. BMC Genomics *9*, 163.

Eisenman, H.C., and Casadevall, A. (2012). Synthesis and assembly of fungal melanin. Appl. Microbiol. Biotechnol. *93*, 931–940.

El-Gebali, S., Mistry, J., Bateman, A., Eddy, S.R., Luciani, A., Potter, S.C., Qureshi, M., Richardson, L.J., Salazar, G.A., Smart, A., et al. (2019). The Pfam protein families database in 2019. Nucleic Acids Res. *47*, D427–D432.

Exophiala dermatitidis (strain ATCC34100/CBS 525.76/NIH/UT8656) (2018) UniProt2.

Exophiala dermatitidis NIH/UT8656 Genome. Assembly NCBI. https://www.ncbi.nlm.nih.gov/assembly/GCF_000230625.1.

Feist, A.M., and Palsson, B. (2008). The growing scope of applications of genome-scale metabolic reconstructions using *Escherichia coli*. Nat. Biotechnol. *26*, 659–667.

Furumura, M., Solano, F., Matsunaga, N., Sakai, C., Spritz, R.A., and Hearing, V.J. (1998). Metal ligand-binding specificities of the tyrosinase-related proteins. Biochem. Biophys. Res. Commun. *242*, 579–585.

García-Borrón, J.C., and Solano, F. (2002). Molecular anatomy of tyrosinase and its related proteins: beyond the histidine-bound metal catalytic center. Pigment Cell Res. *15*, 162–173.

Geis, P.A. (1981). Chemical Composition of the Yeast and Sclerotic Cell Walls of *Wangiella dermatitidis* (University of Texas at Austin).

Geis, P.A., and Szaniszlo, P.J. (1984). Carotenoid pigments of the dematiaceous fungus *Wangiella dermatitidis*. Mycologia *76*, 268–273.

Gudmundsson, S., Agudo, L., and Nogales, J. (2017). Applications of Genome-Scale Metabolic Models of Microalgae and Cyanobacteria in Biotechnology, Microalgae-Based Biofuels and Bioproducts: From Feedstock Cultivation to End-Products (Elsevier Ltd). https://doi.org/10.1016/B978-0-08-101023-5.00004-2.

Gudmundsson, S., and Thiele, I. (2010). Computationally efficient flux variability analysis. BMC Bioinformatics *11*, 2–4.

Ito, S. (2003). IFPCS presidential lecture: a chemist's view of melanogenesis. Pigment Cell Res. *16*, 230–236.

Ito, S., and Wakamatsu, K. (2011). Diversity of human hair pigmentation as studied by chemical analysis of eumelanin and pheomelanin. J. Eur. Acad. Dermatol. Venereol. *25*, 1369–1380.

Kamaraj, B., and Purohit, R. (2014). Mutational analysis of oculocutaneous albinism: a compact review. Biomed. Res. Int. *2014*, 905472.

Khodayari, A., Chowdhury, A., and Maranas, C.D. (2015). Succinate overproduction: a case study of computational strain design using a comprehensive *Escherichia coli* kinetic model. Front. Bioeng. Biotechnol. *2*, 76.

Kumar, J. (2018). Adaptations of Exophiala dermatitidis in Stressful Environments (ETD collection for University of Nebraska - Lincoln), AAI10792984. https://digitalcommons.unl.edu/dissertations/AAI10792984.

Lin, P.C., Saha, R., Zhang, F., and Pakrasi, H.B. (2017). Metabolic engineering of the pentose phosphate pathway for enhanced limonene production in the cyanobacterium Synechocysti s sp. PCC 6803. Sci. Rep. *7*, 17503.

Liu, J., Gao, Q., Xu, N., and Liu, L. (2013). Genome-scale reconstruction and in silico analysis of *Aspergillus terreus* metabolism. Mol. BioSyst. *9*, 1939–1948.

Magnúsdóttir, S., Heinken, A., Kutt, L., Ravcheev, D.A., Bauer, E., Noronha, A., Greenhalgh, K., Jäger, C., Baginska, J., Wilmes, P., et al. (2016). Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota. Nat. Biotechnol. *35*, 81–89.

Mahmood, Alaa, Mohammed, Huda, and Flayyih, May (2015). Purification and Physoichemical Characteristics of Pyomelanin Pigment Produced from Local *Pseudomona aeruginosa* Isolates. World Journal of Pharmaceutical Research *4*, 289–299.

Maranas, C.D., and Zomorrodi, A.R. (2016). Optimization Methods in Metabolic Networks (Wiley).

Moreno, L.F., Vicente, V.A., and de Hoog, S. (2018). Black yeasts in the omics era: Achievements and challenges. Med. Mycol. *56*, 32–41.

Nosanchuk, J.D., and Casadevall, A. (2006). Impact of melanin on microbial virulence and clinical resistance to antimicrobial compounds. Antimicrob. Agents Chemother. *50*, 3519–3528.

Ohkusu, M., Yamaguchi, K., Hata, K., Yoshida, S., Tanaka, R., Nishimura, K., de Hoog, G.S., and Takeo, K. (1999). Cellular and nuclear characteristics of *Exophiala dermatitidis*. Stud. Mycol. *43*, 143–150.

Orth, J.D., Thiele, I., and Palsson, B.O. (2010). What is flux balance analysis? Nat. Biotechnol. *28*, 245–248.

Ou-Yang, H., Stamatas, G., and Kollias, N. (2004). Spectral responses of melanin to ultraviolet a irradiation. J. Invest. Dermatol. *122*, 492–496.

Paolo, W.F., Dadachova, E., Mandal, P., Casadevall, A., Szaniszlo, P.J., and Nosanchuk, J.D. (2006). Effects of disrupting the polyketide synthase gene WdPKS1 in *Wangiella* [*Exophiala*] *dermatitidis* on melanin production and resistance to killing by antifungal compounds, enzymatic degradation, and extreme temperature. BMC Microbiol. *6*, 55.

Papadopoulos, J.S., and Agarwala, R. (2007). COBALT: Constraint-based alignment tool for multiple protein sequences. Bioinformatics *23*, 1073–1079.

Poyntner, C., Blasi, B., Arcalis, E., Mirastschijski, U., Sterflinger, K., and Tafer, H. (2016). The transcriptome of *Exophiala dermatitidis* during ex-vivo skin model infection. Front. Cell Infect. Microbiol. *6*, 136.

Schmaler-Ripcke, J., Sugareva, V., Gebhardt, P., Winkler, R., Kniemeyer, O., Heinekamp, T., and Brakhage, A.A. (2009). Production of pyomelanin, a second type of melanin, via the tyrosine degradation pathway in *Aspergillus fumigatus*. Appl. Environ. Microbiol. *75*, 493–503.

Schoch, C.L., Sung, G.H., López-Giráldez, F., Townsend, J.P., Miadlikowska, J., Hofstetter, V., Robbertse, B., Matheny, P.B., Kauff, F., Wang, Z., et al. (2009). The ascomycota tree of life: a phylum-wide phylogeny clarifies the origin and evolution of fundamental reproductive and ecological traits. Syst. Biol. *58*, 224–239.

Schroeder, W.L., and Saha, R. (2020). OptFill: a tool for infeasible cycle-free gapfilling of stoichiometric metabolic models. iScience *23*, 100783.

Singh, S., Malhotra, A.G., Pandey, A., and Pandey, K.M. (2013). Computational model for pathway reconstruction to unravel the evolutionary significance of melanin synthesis. Bioinformation *9*, 94–100.

Solano, F. (2014). Melanins: skin pigments and much more—types, structural models, biological functions, and formation routes. New J. Sci. *2014*, 1–28.

Spritz, Richard (1994). Molecular genetics of oculocutaneous albinism. Human Molecular Genetics *3* (Review), 1469–1475.

Stolyar, S., Van Dien, S., Hillesland, K.L., Pinel, N., Lie, T.J., Leigh, J.A., and Stahl, D.A. (2007). Metabolic modeling of a mutualistic microbial community. Mol. Syst. Biol. *3*, 92.

Strobel, I., Breitenbach, J., Scheckhuber, C.Q., Osiewacz, H.D., and Sandmann, G. (2009). Carotenoids and carotenogenic genes in *Podospora anserina*: engineering of the carotenoid composition extends the life span of the mycelium. Curr. Genet. *55*, 175–184.

Sudhadham, M., Prakitsin, S., Sivichai, S., Chaiyarat, R., Dorrestein, G.M., Menken, S.B., and de Hoog, G.S. (2008). The neurotropic black yeast *Exophiala dermatitidis* has a possible origin in the tropical rain forest. Stud. Mycol. *61*, 145–155.

Szaniszlo, P.J. (2002). Molecular genetic studies of the model dematiaceous pathogen *Wangiella dermatitidis*. Int. J. Med. Microbiol. *292*, 381–390.

Thiele, I., and Palsson, B.Ø. (2010). A protocol for generating a high-quality genome-scale metabolic reconstruction. Nat. Protoc. *5*, 93–121.

Toledo, Andrea Vanesa, Franco, Mario Emilio Ernesto, Lopez, Silvina Marianela Yanil, Troncozo, Maria Ines, Saparrat, Mario Carlos Nazareno, and Balatti, Pedro Alberto (2017). Melanins in fungi: types, localization and putative biological roles. Physiol. Mol. Plant Pathol. *99*, 2–6.

Upadhyay, S., Xu, X., Lowry, D., Jackson, J.C., Roberson, R.W., and Lin, X. (2016). Subcellular compartmentalization and trafficking of the biosynthetic machinery for fungal melanin. Cell Rep. *14*, 2511–2518.

Vongsangnak, W., Olsen, P., Hansen, K., Krogsgaard, S., and Nielsen, J. (2008). Improved annotation through genome-scale metabolic modeling of *Aspergillus oryzae*. BMC Genomics *9*, 245.

Williams, T.C.R., Poolman, M.G., Howden, A.J., Schwarzlander, M., Fell, D.A., Ratcliffe, R.G., and Sweetlove, L.J. (2010). A genome-scale metabolic model accurately predicts fluxes in central carbon metabolism under stress conditions. Plant Physiol. *154*, 311–323.

Yamamoto, H.Y., and Bangham, A.D. (1978). Carotenoid organization in membranes. Thermal transition and spectral properties of carotenoid-containing liposomes. Biochim. Biophys. Acta *507*, 119–127.

Yizhak, K., Chaneton, B., Gottlieb, E., and Ruppin, E. (2015). Modeling cancer metabolism on a genome scale. Mol. Syst. Biol. *11*, 817.

Yu, C., and Lin, C. (2004). Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on n -peptide compositions. Protein Sci. *13*, 1402–1406.

Zaghdoudi, K., Ngomo, O., Vanderesse, R., Arnoux, P., Myrzakhmetov, B., Frochot, C., and Guiavarc'h, Y. (2017). Extraction, identification and photo-physical characterization of Persimmon (Diospyros kaki L.) carotenoids. Foods *6*, 4.

Zhang, X., Tervo, C.J., and Reed, J.L. (2016). Metabolic assessment of E. coli as a Biofactory for commercial products. Metab. Eng. *35*, 64–74.

iScience, Volume *23*

Supplemental Information

Computation-Driven Analysis of Model Polyextremotolerant Fungus *Exophiala dermatitidis*: Defensive Pigment Metabolic Costs and Human Applications

Wheaton L. Schroeder, Steven D. Harris, and Rajib Saha

SUPPLEMENTAL VIDEO, DATA, AND TABLES TITLES AND LEGENDS

Figure S1: Carotenoid synthesis pathway, part 1, Related to Figure 3. This figure show the synthesis pathway of carotenoids from acetyl-CoA through geranylgeranyl diphosphate. Included in this figure are chemical structures of the primary metabolites in the pathway; cofactors and enzymes used in each step; and shadow prices under the investigated carbon-limited conditions.

Figure S2: Carotenoid synthesis pathway, part 2, Related to Figure 3. This figure show the synthesis pathway of carotenoids from geranylgeranyl diphosphate through to carotenoid intermediates. Included in this figure are chemical structures of the primary metabolites in the pathway; cofactors and enzymes used in each step; and shadow prices under the investigated carbon-limited conditions.

Figure S3: Carotenoid synthesis pathway, part 3, Related to Figure 3. This figure show the synthesis pathway of carotenoids from geranylgeranyl diphosphate through to carotenoid products. Included in this figure are chemical structures of the primary metabolites in the pathway; cofactors and enzymes used in each step; and shadow prices under the investigated carbon-limited conditions.

Data S1: Hidden Markov Model Analysis of *E. dermatitidis* tyrosinase gene copies, Related to Figure 5. This file contains the basic information (screenshot of results and very brief discussion) of the Hidden Markov Model analysis of each *E. dermatitidis* tyrosinase gene copy of tyrosinase using the Pfam tool.

TRANSPARENT METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Software and Algorithms		
Generalized Algebraic Modeling System (GAMS) version 24.7.4	GAMS Development Corp. (URL: www.gams.com/)	N/A
CPLEX Solver version 12.6	GAMS Development Corp. (URL: www.gams.com/)	N/A
Perl version 5.26 (for Unix)	Perl.org (URL: www.perl.org)	N/A
Strawberry Perl version 5.24.0.1 (for Windows)	Perl.org (URL: http://strawberryperl.com/)	N/A
The world-wide-web library for Perl, module 6.39	Perl meta::cpan (URL: metacpan.org/pod/LWP)	N/A
Python version 3.3 (for Unix)	Python Software Foundation (URL: www.python.org)	N/A
Kyoto Encyclopedia of Genes and Genomes (KEGG)	Kanehisa Laboratories (URL: www.kegg.jp)	RRID:SCR_012773
Exophiala dermatitidis (NIH/UT8656) annotated genome.	UniProt (URL: www.uniprot.org)	N/A
Exophiala dermatitidis (NIH/UT8656) annotated genome.	National Center for Biotechnology Information (URL: www.ncbi.nlm.nih.gov)	N/A
Other		
Holland Computing Center, Crane Cluster (64 GB RAM, Intel Xenon E5-2670 2.60 GHz processor, 2 CPUs per node)	Holland Computing Center, University of Nebraska (URL: hcc.unl.edu/)	N/A

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Rajib Saha (rsaha2@unl.edu). No materials were generated in this work including cells, DNA, antibodies, reagents, organisms, mouse strains, or ES cells.

Materials Availability Statement

This study has produced several unique software codes in the form of GAMS, Perl, or Python programming languages/tools. These are included in the GitHub (DOI: 10.5281/zenodo.3608172) repository which accompany this work.

METHOD DETAILS

Use of E. dermatitidis genome annotation information. Genome annotations of *E. dermatitidis* were retrieved from NCBI (*Exophiala dermatitidis NIH/UT8656 Genome Assembly*, 2011) and UniProt (*Exophiala dermatitidis (strain ATCC34100/CBS 525.76/NIH/UT8656)*, 2018) databases. These initial retrievals contained 9562 and 9391 Open Reading Frames (ORFs), respectively, with up to 4.9% of ORFs labeled with Enzyme Classification (EC) numbers. As EC numbers are often used to establish Gene-Protein-Reaction (GPR) links in a GSM, programming scripts and a library of programming functions were devised to automatically search the BRENDA database with the protein name to attempt to discover EC numbers for as many proteins as possible. These can be found in the GitHub "E_dermatitidis_model" repository which accompanies this work. This resulted in 2020 (21.5%) and 1724 (18.0%) of UniProt and NCBI ORFs, respectively, that

correspond to at least one specific EC number. Accepting only single EC number BRENDA search results (to discount matches due to ambiguous names), and only those which are present in both annotations resulted in 532 EC numbers. Using these EC numbers, the reactions which KEGG indicated could be catalyzed by these EC numbers (determined by File SFF) were used to form the first draft model. These reactions were assigned to subcellular compartments by inputting the FASTA corresponding to the 537 EC numbers to the CELLO predictor for subcellular localization (Yu and Lin, 2004)(Yu, Lin and Hwang, 2006). These results, included in Supplemental File 4, predicted 533 cytosolic, 435 mitochondrial, 66 extracellular, 6 lysosomic, 117 peroxisomic, 144 nucleic, and 5 endoplasmic reticulate reactions. Cytosolic, mitochondrial, and extracellular reactions were selected for incorporation to the model. Peroxisomic reactions were not included due to large number of metabolic gaps, resulting in many reactions which produced and/or consumed metabolites not present elsewhere in *Exophiala* metabolism, and for which literature evidence justifying their inclusion could not be found. Further, the lack of information in literature as to the metabolism which occurs in fungal peroxisomes and metabolite transporters further hinders accurate reconstruction of peroxisome metabolism to the extent that accurate reconstruction may not be possible at present. Nucleic reactions were not included as most reactions in this organelle involve the synthesis, breakdown, modification, or maintenance of RNA and DNA, and do not generally participate in other metabolic processes. Further, as with the peroxisome, some metabolites were present here which were not present elsewhere in *Exophiala* metabolism (other than DNA/RNA). The mitochondrial compartment was separated to inner and outer mitochondria, resulting in four distinct model compartments. The outer mitochondrion is modeled as compartment to store protons pumped by oxidative phosphorylation and another biologically relevant membrane across which transport must occur. The set of cytosolic, mitochondrial, and extracellular reactions were used in the definition of the first draft model of *Exophiala*.

Choice of Aspergillus models for metabolic gapfilling. Utilizing the most comprehensive phylogenetic tree for the Ascomycota phylum found (Schoch *et al.*, 2009), the phylogenetic branches of the Ascomycota phylum where investigated for Genome-Scale Models (GSMs) which were created for related organisms (nearest branches were investigated first). The most closely related models identified belong to the *Aspergillus* genus, namely: *A. niger* (Andersen, Nielsen and Nielsen, 2008), *A. nidulans* (David *et al.*, 2008), *A. oryzae* (Vongsangnak *et al.*, 2008), and *A. terreus* (Liu *et al.*, 2013) which all belong to the same class as *E. dermatitidis*, Eurotiomycetes. No other genome-scale models belonging to this same class were identified. Genome-Scale models belonging to the same phylum as *E. dermatitidis* were considered for inclusion in the definition of automated gapfilling database, but this was dismissed for multiple reasons. First, this would result in a linear increase with the number of included models in the number of OptFill runs needed based on commonality of enzymes, resulting in a less tractable study. Second, this would reduce the number of core enzymes to those common to the phylum as opposed to the class. Third, this allows for a more conservative metabolic reconstruction, reducing the chances of adding false functionalities. Fourth, manually completed BLASTp analyses between *E. dermatitidis* and the model Ascomycete, *Saccharomyces cerevisiae*, showed poor alignments between sequences encoding the same enzyme. Fifth, there are acknowledged conserved homologs between *Aspergillus* and *Exophiala* species including two tyrosinase enzymes, polyketide synthase, and alpha-beta hydrolase (Chen *et al.*, 2014). Finally, *Aspergillus* genomes are larger, 34 Mbp (*A. niger*), 30 Mbp (*A. nidulans*), 37.8 Mbp (*A. oryzae*), and 29.4 Mbp (*A. terreus*) (*National Center*

for Biotechnology Information, no date), than the *E. dermatitidis* genome, 26.4 Mbp (*National Center for Biotechnology Information*, no date). The larger genome of *Aspergillus* species would likely encode for more metabolic functionalities that *E. dermatitidis*. However, model Ascomycota species likely encode for fewer metabolic functionalities in that their genomes are significantly smaller than *E. dermatitidis*, 11.8 Mbp for *S. cerevisiae* and 20.2 Mbp for *Y. lipolitica*. For these reasons, it was deemed appropriate to restrict the databases associated with filling metabolic gaps to functionalities identifiable in *Aspergillus* species.

Consensus of Aspergillus models. The relatively small percentages of ORFs with assigned EC numbers by annotations or by information from BRENDA suggested the need to explore GSMs of related species to identify core enzymes. The *Aspergillus* genus was identified as closely related (Schoch *et al.*, 2009) with four GSM models: *A. niger* (Andersen, Nielsen and Nielsen, 2008), *A. nidulans* (David *et al.*, 2008), *A. oryzae* (Vongsangnak *et al.*, 2008), and *A. terreus* (Liu *et al.*, 2013). Models of *Saccharomyces cerevisiae* were not used at this stage of the curation process as *S. cerevisiae* is phylogenetically much more distant to *E. dermatitidis* than are *Aspergillus* species (Schoch *et al.*, 2009). Using the GPR links in the *Aspergillus* models, provided in the form of EC numbers, the links were sorted into bins of full consensus, common to three, common to two, and only present in one. EC numbers in the full consensus bin were added to the *Exophiala* model by repeating the process used on the *Exophiala* EC numbers and using the compartmentalization from the *Aspergillus* models. The other bins were similarly converted to lists of reactions with compartmentalization from the *Aspergillus* models. These three reaction lists are the databases used in the application of OptFill (Schroeder and Saha, 2019) to the *Exophiala* model. Further, transport reactions were taken from the *Aspergillus* models and added to the *Exophiala* model as needed.

Bidirectional BLASTp of Aspergillus consensus enzymes. In order to create GPR links between *Aspergillus* enzymes and reactions added to the *Exophiala* models, a program which performs a bidirectional BLASTp on a list of enzymes subject to certain constraints and with the intent of identifying a gene encoding each enzyme in the list was created, and is included in GitHub "E_dermatitidis_model" repository. The constraint file specifies the target organism (in this work, *Exophiala*), the file path in which FASTAs and BLASTp results will be deposited to, the expect value upper bound cut-off (for which a match to be accepted), the percent positive substitution lower bound value (for matches in which the percent positive substitution value being greater than or equal to the cut off), and related species from which to take reference sequences for an enzyme. The Bidirectional Blast Program (hereafter BBP) begins by reading the constraints and enzyme list files. The workflow followed by the BBP is shown in Supplemental Figure SDD. In short, BBP takes the enzyme classification (EC) number and uses it to look up the amino acid sequences for genes encoding that EC number in the related organism. The amino acid sequences (in the form of a FASTA file) of the related organism (in this work, *A. nidulans*, *A. niger*, *A. terreus* or *A. oryzae*) is BLASTed against the target organism (in this work, *Exophiala*). This is called the "forward BLAST". Should this forward BLAST result in a significant match according to the cutoffs specified in the constraints file, the amino acid sequence from *Exophiala* is then BLASTed against the related organism from which the EC producing amino acid sequence is taken. This is referred to as the "backward BLAST". Should the backward BLAST results be significant according to the cutoffs in the constraint file, the match between the two sequences is accepted. A summary of the

BLAST results can be found in Supplemental File 2, FASTAs for the sequences can be found in the GitHub "E_dermatitidis_model" repository.

Definition of biomass composition. The definition of biomass is provided in supplemental File 4. Literature evidence was sought out for the biomass composition of *Exophiala*, beginning with the composition of the cell wall. The cell wall composition can be found in Table 2 and is from data for *Exophiala* grown at $37^{o}\ C$ (Geis, 1981). As this work did not distinguish between types of melanin in the cell wall of *Exophiala*, and no follow-up study was found which made this distinction, each of the three melanins, namely DNH melanin, eumelanin, and pyomelanin, is assumed to contribute equally to the cell wall mass. Further, the composition of the lipid term was unspecified. No data was found as to the lipid composition of *Exophiala*, and therefore the lipid composition of *A. terreus* grown using glucose as a carbon source at $28^{0}\ C$ was used (Kumar and Vatsyayan, 2010). Only lipids with KEGG identifiers were included in the lipid composition definition, accounting for 82.3% of the lipid composition of *A. terreus* (by weight percentage) (Kumar and Vatsyayan, 2010). In addition, there was no information on the fraction of *Exophiala* cell mass that was accounted for by the cell wall itself; however, *S. cerevisiae* cell walls account for 15-30% of the total cell mass (Lipke and Ovalle, 1998). For *Exophiala* , it was assumed that 25% of the cell mass is cell wall as the cell wall of this species has been described as "thick" (Chen *et al.*, 2014)(Kumar and Vatsyayan, 2010)(Schnitzler *et al.*, 1999). The next set of literature evidence sought for biomass composition is the cell without the cell wall. No literature evidence was found, and so to be consistent with the lipid composition information and considering that the *A. terreus* model is the most recently published of the four *Aspergillus* models, its biomass composition was used for the cell. Finally, the carotenoid contribution to biomass was considered. Again, lack of information pertaining to *Exophiala* led to using carotenoid biomass composition data from another organism, in this case, *Podospora anserina*, for which approximately 3.47% of cell mass is carotenoids (Strobel *et al.*, 2009). Both *P. anserina* and *Exophiala* are Ascomycota, but phylogenetically diverge at the class level. Unfortunately, no similar data was able to be found for a species which was phylogenetically closer to *Exophiala*. The remaining cell weight, 71.53%, was assumed to be composed of the cell membrane and all biomass components enclosed within it (such as proteins and lipids). As the ratios of carotenoids, cell wall, and other cell components were determined to be important, biomass composition was divided into cell wall, cell, and carotenoid pseudometabolites. These pseudometabolites represented the mass contributions of fixed stoichiometeries of metabolites which comprise that portion of biomass. Each of these then had their own pseudo-molecular weight. These three pseudometabolites were then combined in a pseudoreaction to form biomass. The stoichiometric ratio of these pseudometabolites in the biomass reaction was determined by using the Solver tool in Microsoft Excel whose objective was a biomass molecular weight of 1000 mg/gDW·h. Ratios between pseudometabolites were enforced in this analysis to preserve ratios as described above. The resulting biomass composition can be found in Supplemental File 4.

Creation of first and second drafts of E. dermatitidis GSM. The first draft of the *Exophiala* GSM was the combination of the set of reactions which exist in *Exophiala* as determined by the analysis of the annotated genomes, the analysis of consensus *Aspergillus* enzymes, and the defined biomass composition. This first draft model did not produce biomass, melanins, or carotenoids. Through manual curation and addition of reactions related to melanin and carotenoid synthesis, both classes of pigments were produced by the second *Exophiala* draft model. Thermodynamically Infeasible

Cycle (TICs) in the draft model were manually addressed. Further, some reactions were needed to be manually added to the model to ensure biomass production. These reactions were taken from a GSM of *S. cerevisiae*, *i*Sce926 (Chowdhury, Chowdhury and Maranas, 2015), and a list of reactions derived from the common to four *Aspergillus* models enzyme overlap (Andersen, Nielsen and Nielsen, 2008)(David *et al.*, 2008)(Vongsangnak *et al.*, 2008)(Liu *et al.*, 2013) using code included in the GitHub "E_dermatitidis_model" repository. Reactions from the common to three *Aspergillus* list which were used in manual curation were removed from that list before performing OptFill on the second draft model. Notes related to the curation process can be found in the GitHub "E_dermatitidis_model" repository. Once TICs were eliminated, the model could produce all pigment molecules and biomass and utilize multiple literature-supported carbon sources including sucrose, ethanol, acetate, and glucose. The second draft *Exophiala* model was then considered complete. This second draft model still had a significant number of metabolic gaps, particularly in secondary metabolism, as evidenced by Flux Variability Analysis (FVA) (Gudmundsson and Thiele, 2010). When applied to this model, FVA showed that only 711 of 1587 reaction present in the model were capable of carrying flux (about 44.8%). This model could produce 591 metabolites (of a total of 1839). The maximum rate of growth of this model was 0.0952 h^{-1}.

Update to the OptFill algorithm. In the process of applying OptFill to draft models of *Exophiala*, which are the largest database/model pairs to which OptFill has thus far been applied (Gudmundsson and Thiele, 2010), it was discovered that additional constraints were necessary in order that the algorithm is not sensitive to solver options used. Specifically, these constraints were required in the Connecting Problems (CPs) of OptFill, and are listed below in equations (13) through (17), along with the full formulation for the first CP. The second and third CPs used here are related to the first CP in the same manner as detailed in Schroeder and Saha (2020).

$$maximize\ Z_{met} = \sum_{i_m \in I^M} x_{i_m} \tag{1}$$

Subject to

$$\sum_{j_{db} \in J^{DB}} \zeta_{j_{db}} \geq 1 \tag{2}$$

$$\rho_{j_{db}} v_{j_{db}}^{LB} \leq v_{j_{db}} \leq \delta_{j_{db}} v_{j_{db}}^{UB} \quad \forall j_{db} \in J^{DB} \tag{3}$$

$$\theta_j v_j^{LB} \leq v_j \leq (1-\theta_j) v_j^{UB} - \epsilon\theta_j \quad \forall j \in J \tag{4}$$

$$(1-\lambda_j) v_j^{LB} + \epsilon\lambda_j \leq v_j \leq \lambda_j v_j^{UB} \tag{5}$$

$$x_i \leq \sum_{j \in J} [\lambda_j \xi_{i,i} + \theta_j \psi_{i,j}] \quad \forall i \in I \tag{6}$$

$$x_b = 1 \quad \forall b \in B \subset I \tag{7}$$

$$\sum_{j \in J} S_{ij} v_j = 0 \quad \forall i \in I \tag{8}$$

$$\zeta_{j_{db}} \leq \sum_{i \in I} [\lambda_j \xi_{i,i} + \theta_j \psi_{i,j}] \quad \forall j_{db} \in J^{DB} \tag{9}$$

$$\delta_{j_{db}} + \rho_{j_{db}} - \omega_{j_{db}} = \zeta_{j_{db}} \quad \forall j_{db} \in J^{DB} \tag{10}$$

$$\omega_{j_{db}} \leq \delta_{j_{db}} \quad \forall j_{db} \in J^{DB} \tag{11}$$

$$\omega_{j_{db}} \leq \rho_{j_{db}} \quad \forall j_{db} \in J^{DB} \tag{12}$$

$$\omega_{j_{db}} \leq \theta_{j_{db}} + \lambda_{j_{db}} \quad \forall j_{db} \in J^{DB} \tag{13}$$

$$\rho_{j_{db}} \leq \theta_{j_{db}} + \lambda_{j_{db}} \quad \forall j_{db} \in J^{DB} \tag{14}$$

$$\delta_{j_{db}} \leq \theta_{j_{db}} + \lambda_{j_{db}} \quad \forall j_{db} \in J^{DB} \tag{15}$$

$$\delta_{j_{db}} \geq v_{j_{db}} \qquad \forall j_{db} \in J^{DB} \qquad (16)$$

$$\rho_{j_{db}} \geq v_{j_{db}} \qquad \forall j_{db} \in J^{DB} \qquad (17)$$

$$\sum_{j_{db} \in J^{DB}} \delta_{j_{db}} \left(\delta'_{s_c, j_{db}}\right) \leq \sum_{j_{db} \in J^{DB}} \delta'_{s_c, j_{db}} - \sigma_{s_c} \qquad \forall s_c \in S_c \qquad (18)$$

$$\sum_{j_{db} \in J^{DB}} \rho_{j_{db}} \left(\rho'_{s_c, j_{db}}\right) \leq \sum_{j_{db} \in J^{DB}} \rho'_{s_c, j_{db}} - \left(1 - \sigma_{s_c}\right) \qquad \forall s_c \in S_c \qquad (19)$$

$$\sum_{j_{db} \in J^{DB}} \left(\delta'_{s_c, j_{db}} - \delta_{j_{db}}\right) + \sum_{j_{db} \in J^{DB}} \left(\rho'_{s_c, j_{db}} - \rho_{j_{db}}\right) \geq \left(\sum_{j_{db} \in J^{DB}} \omega'_{s_c, j_{db}}\right) + 1 \qquad \forall s_c \in S_c \qquad (20)$$

$$\sum_{j_{db} \in J^{DB}} \delta_{j_{db}} \left(\alpha'_{s_f, j_{db}}\right) \leq \sum_{j_{db} \in J^{DB}} \alpha'_{s_f, j_{db}} - \tau_{s_f} \qquad \forall s_f \in S_f \qquad (21)$$

$$\sum_{j_{db} \in J^{DB}} \rho_{j_{db}} \left(\beta'_{s_f, j_{db}}\right) \leq \sum_{j_{db} \in J^{DB}} \beta'_{s_f, j} - \left(1 - \tau_{s_f}\right) \qquad \forall s_f \in S_f \qquad (22)$$

Where symbols used are defined as follows.

$Fixed\ Values\ Unique\ to\ CP1$

$M = 1E3 \equiv a\ very\ large\ number$

$$\delta'_{s_c, j_{db}} = \begin{cases} 1\ if\ reaction\ j_{db}\ is\ added\ in\ the\ forward\ direction\ from\ the \\ database\ in\ solution\ s_c \\ 0\ otherwise \end{cases}$$

$$\rho'_{s_c, j_{db}} = \begin{cases} 1\ if\ reaction\ j_{db}\ is\ added\ in\ the\ backward\ direction\ from\ the \\ database\ in\ solution\ s_c \\ 0\ otherwise \end{cases}$$

$$\omega'_{s_o, j_{db}} = \begin{cases} 1\ if\ reaction\ j_{db}\ is\ added\ in\ the\ forward\ direction\ from\ the \\ database\ in\ solution\ s_o \\ 0\ otherwise \end{cases}$$

$$\xi_{i,j} = \begin{cases} 1\ if\ metabolite\ i\ is\ on\ the\ RHS\ of\ reaction\ j\ \left(S_{i,j} > 0\right) \\ 0\ otherwise \end{cases}$$

$$\psi_{i,j} = \begin{cases} 1\ if\ metabolite\ i\ is\ on\ the\ LHS\ of\ reaction\ j\ \left(S_{i,j} < 0\right) \\ 0\ otherwise \end{cases}$$

$Variables\ Unique\ to\ CP1$

$$\delta_{j_{db}} = \begin{cases} 1\ if\ reaction\ j_{db}\ is\ added\ in\ the\ forward\ direction\ from\ the\ database \\ 0\ otherwise \end{cases}$$

$$\rho_{j_{db}} = \begin{cases} 1\ if\ reaction\ j_{db}\ is\ added\ in\ the\ backwards\ direction\ from\ the\ database \\ 0\ otherwise \end{cases}$$

$$\omega_{j_{db}} = \begin{cases} 1\ if\ reaction\ j_{db}\ is\ added\ reversibly\ from\ the\ database\ \left(\delta_{j_{db}} = \rho_{j_{db}} = 1\right) \\ 0\ otherwise \end{cases}$$

$$\zeta_{j_{db}} = \begin{cases} 1\ if\ reaction\ j_{db}\ is\ part\ of\ the\ solution \\ 0\ otherwise \end{cases}$$

$$\theta_j = \begin{cases} 1\ if\ reaction\ is\ proceding\ in\ backwards\ direction\ \left(v_j < 0\right) \\ 0\ otherwise \end{cases}$$

$$\lambda_j = \begin{cases} 1\ if\ reaction\ is\ proceding\ in\ forwards\ direction\ \left(v_j > 0\right) \\ 0\ otherwise \end{cases}$$

$$x_i = \begin{cases} 1\ if\ connected\ model\ produces\ metabolite\ i \\ 0\ otherwise \end{cases}$$

$\sigma_{s_c} \in [0,1] \equiv binary\ variable\ which\ ensures\ that\ the\ solution\ is\ unique\ from\ previous\ solutions\ in\ at\ least\ one\ direction\ of\ one\ database\ reaction$

$\tau_{s_c} \in [0,1] \equiv binary\ variable\ which\ ensures\ that\ the\ solution\ is\ free\ from\ TICs\ in\ that\ at\ least\ one\ direction\ of\ one\ database\ reactions\ which\ could\ cause\ a\ TIC\ is\ not\ added\ in\ the\ TIC-causing\ direction\ for\ each\ TIC\ identified\ by\ the\ TFP$

Equations displayed above, with the exception of equations (13) through (17), are identical to the original formulation of OptFill. In short, additional constraints (13), (14), and (15) explicitly link binary variables noting the reaction direction and binary variables relating the direction of database reactions which are added to the model. This reduces the impact of any feasibility relaxation assumptions made by the solver in attempting to solve the CPs. Additional constraints (16) and (17) restrict the range of each reaction rate, v_j, to be less than one (as flux magnitude is not important and this allows effectively tighter optimization criteria), while at the same time requiring $\delta_{j_{db}}$ and $\rho_{j_{db}}$ to have a non-zero value if the metabolic flux through that database reaction is non-zero. Theoretically, this addressed by constraint equation (3), but these statements again reduce the effect of feasibility relaxations.

First use of OptFill to address metabolic gaps. To address the metabolic gaps in the second draft model, OptFill (Gudmundsson and Thiele, 2010) was used first with a list of reactions derived from the list of enzymes common to four *Aspergillus* models. The database was reduced to a manageable size (e.g. one that would allow at least one solution to the CPs in less than one week) over six rounds of identifying TICs and pruning the database of reactions which caused the most TICs. By the end of this pruning, the database consisted of 241 reactions, had 82 potential TICs with the model (largest size 12 reactions), and two connecting problem solutions. The first solution, which could produce 620 metabolites by adding 20 reversible reactions, was accepted over the second solution, which could produce 619 metabolites by adding seven reversible and 11 irreversible reactions. Adding the first CPs solution to the second draft model produced the third draft model consisting of the 1607 reactions, of which 749 are capable of carrying flux (about 46.6%). The maximum rate of growth of this model was 0.0989 h^{-1} by allowing for up to 10 $mmol \cdot gDW^{-1} \cdot h^{-1}$ uptake of one of ethanol, sucrose, glucose, and acetate along with sufficient amount of nitrate, sulfate, and phosphate.

Second and third use of OptFill to address metabolic gaps. This process was repeated twice: the first time using the list of reactions derived from the list of enzymes common to two *Aspergillus* models, and the second time using the list of reactions derived from the list of enzymes common to two *Aspergillus* models. For more details on the results of each step, see Supplemental File 5. The end result is the final model, *i*Ede2091, consisting of 1630 reactions, of which 793 are capable of carrying flux (48.7%). The maximum rate of growth of this model was 0.0989 h^{-1} by allowing for 10 $mmol \cdot gDW^{-1} \cdot h^{-1}$ uptake of ethanol, sucrose, glucose, acetate, nitrate, sulfate, and phosphate. In this growth condition, carbon is the limiting nutrient. It should be noted that, on a minimal media where sucrose is provided as the primary carbon source, that is at a concentration two orders of magnitude higher than any other potential carbon source, the growth rate of *Exophiala* is approximately 0.105 h^{-1} (Dadachova *et al.*, 2007); however, since no rate measures were taken in the indicated study, it is difficult to interpret the accuracy of the modeled growth rate of *i*Ede2091.

Bidirectional BLASTp to investigate OptFill solution viability. For each OptFill solution incorporated into the draft models, a bidirectional BLASTp analysis was performed on enzymes linked to the reactions in each OptFill solution. For the first OptFill solution, containing reactions linked to enzymes common to three of four *Aspergillus* models, 20 enzymes were identified as linked to this set of reactions. Using the same bidirectional BLASTp procedure as previously described, 11 of these enzymes were identified in the *E. dermatitidis* genome, being matched to

21 genes. These genes were all annotated in the *E. dermatitidis* genome; therefore these enzymes may not have been identified by the BRENDA search of *E. dermatitidis* enzyme annotations due to sensitivity of the algorithm used for this search. These matches give a genetic basis for the inclusion of 11 of these reactions, in addition to the evidence that all these enzymes are supported in phylogenetically related organisms.

For the second OptFill solution, containing 3 reactions, 3 enzymes were identified as linked to the set of reactions in the solution, and two of these enzymes where identified in the *E. dermatitidis* genome. These two enzymes were linked to two genes. These genes were all annotated in the *E. dermatitidis* genome; therefore these enzymes may not have been identified by the BRENDA search of *E. dermatitidis* enzyme annotations due to sensitivity of the algorithm used for this search to the annotated string. These matches give a genetic basis for the inclusion of 2 of these reactions, in addition to the evidence that all these enzymes are supported in phylogenetically related organisms.

For the third OptFill solution, containing 21 reactions, 17 enzymes were identified as linked to the set of reactions in the solution, and 8 of these enzymes where identified in the *E. dermatitidis* genome. These 8 enzymes were linked to 18 genes. These genes were all annotated in the *E. dermatitidis* genome; therefore these enzymes may not have been identified by the BRENDA search of *E. dermatitidis* enzyme annotations due to sensitivity of the algorithm used for this search to the annotated string. These matches give a genetic basis for the inclusion of 13 of these reactions, in addition to the evidence that all these enzymes are supported in phylogenetically related organisms.

Flux Balance and shadow price analyses. Flux Balance Analysis (FBA) is a tool to study distribution of fluxes subject to an objective function (often growth) and certain constraints (e.g., mass balance and nutrient availability) for an underdetermined network (Orth, Thiele and Palsson, 2010), and was performed as previously described (Orth, Thiele and Palsson, 2010)(Gianchandani, Chavali and Papin, 2010). The dual formulation of FBA, and the definition of shadow price, was derived as described by Zomorrodi and Costas (Maranas and Zomorrodi, 2016). The shadow price of a metabolite is the change in the value of the objective function used in FBA (growth) that would result from producing one more unit (mmol·gDW^{-1}·h^{-1}) of that metabolite. The shadow price is the λ_i variable of the dual formulation of the FBA problem shown below. The shadow price relating to all 36 growth conditions was calculated using the primary and dual formulations shown below. The primal FBA problem is as follows.

$$maximize\ z_{prime} = v_{biomass} \tag{23}$$

Subject to:

$$\sum_{j \in J} S_{ij} v_j = 0 \qquad \forall i \in I \tag{24}$$

$$v_j^{LB} \leq v_j \leq v_j^{UB} \qquad \forall j \in J \tag{25}$$

The dual FBA problem is as follows.

$$maximize\ z_{dual} = \sum_{j \in J} v_j^{UB} \mu_j^{UB} + \sum_{j \in J} v_j^{LB} \mu_j^{LB} \quad (26)$$

Subject to:

$$\sum_{i \in I} S_{ij} \lambda_i - \mu_j^{LB} + \mu_j^{UB} = 0 \qquad \forall j \in J - biomass \quad (27)$$

$$\sum_{i \in I} S_{ij} \lambda_i - \mu_j^{LB} + \mu_j^{UB} = 1 \qquad \forall j \in biomass \quad (28)$$

By applying strong duality theory using the following constraint, both primal and dual variables may be explicitly solved.

$$z_{dual} = z_{primal} \quad (29)$$

In the reconstruction of the *i*Ede2091 model, it was noted that the availability of three nutrient atoms, namely carbon, nitrogen, and sulfur, could limit the growth of the *i*Ede2091 model. Further the model can grow on four different carbon sources. Tt was decided to investigate the effect of growth limiting nutrients on the shadow price of defensive pigments and their precursors under 36 unique growth conditions, where each carbon source/limiting atom pair is investigated under low, moderate, and high availability.

Identification of E. dermatitidis tyrosinase enzymes and attempted identification of tyrosinase related proteins. Four tyrosinase enzymes have been annotated in the *Exophiala* genome and noted in literature (Chen *et al.*, 2014). In order to ensure that all gene copies of *Exophiala* tyrosinase were accounted for, the four tyrosinase genes from *E. dermatitidis* were BLASTed against the *Exophiala* genome using non-redundant BLASTp. No accessions which were not previously annotated as tyrosinase were identified, see the GitHub "E_dermatitidis_model" repository for the BLAST results related to tyrosinase. As tyrosinase-related proteins share high sequence similarity to tyrosinases of a species (Furumura *et al.*, 1998), all four tyrosinase sequences for *Exophiala* were BLASTed against its own genome, again using non-redundant BLASTp. No significant matches were found except for known tyrosinases. As there is no literature evidence for *Exophiala* or *Aspergillus* species with tyrosinase-related proteins, it was concluded from this that no tyrosinase-related proteins are encoded for by the *Exophiala* genome.

Comparison of E. dermatitidis tyrosinase gene copies to Hidden Markov Model (HMM) tyrosinase sequences. A comparison of each tyrosinase sequence in *E. dermatitidis* was made to the Hidden Markov Model tyrosinase sequence using the Pfam tool (El-Gebali *et al.*, 2019). The amino acid sequence for each *E. dermatitidis* gene copy was used as the search sequence. All gene copies had strong sequence alignments to the tyrosinase HMM, with gene copies unique to *Exophiala* matching weakly to the tyrosinase C HMM as well. All sequence alignments had expect values between 2.0E-38 and 3.1E-54, showing very strong agreement.

Comparison of E. dermatitidis tyrosinase gene copies to human tyrosinase alleles. First, the amino acid sequences of *Exophiala* tyrosinases were BLASTed against the human genome using non-redundant BLASTp. This produced no significant matches (see the tyrosinase sequence alignments provided in the GitHub "E_dermatitidis_model" repository) initially suggesting that

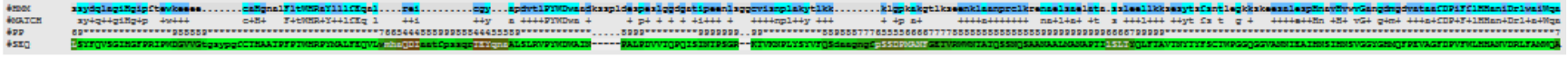
these enzymes were quite different. However, a COBALT amino acid sequence alignment was performed comparing three human alleles for tyrosinase, tyrosinase-related protein sequences from human, and the tyrosinase reference sequences for *Exophiala*. For the human alleles chosen, one was a reference sequence, one an albino sequence for oculocutaneous albinism A1, and one sequence from an individual of the Bantu peoples of Kenya (Hudjashov, Villems and Kivisild, 2013), representing a population susceptible to the ill-effects of albinism (Brilliant, 2015). An independent COBALT sequence alignment was also performed with only the three human alleles selected to identify the sequential differences between the three alleles. Both COBALT alignments are provided in the GitHub "E_dermatitidis_model" repository, and a visualization of the results is provided in Figure 5, with particular attention paid to the active site of tyrosinase which is the two copper-binding domains. Visualization highlighting uses the 3-bit conservation score setting was used for highlighting sequence similarities as it seems a moderately-strict setting and no standard for this highlighting scheme was identified in literature. Literature was used to identify both tyrosinase active sites, CuA, from approximately residues 173 to 220 in human tyrosinase (Furumura *et al.*, 1998)(García-Borrón and Solano, 2002), and CuB, from approximately residues 361 to 403 in human tyrosinase (Furumura *et al.*, 1998)(García-Borrón and Solano, 2002)(Spritz *et al.*, 1997). Labels for the significance of highly conserved residues are taken from the analysis of García-Borrón and Solano (2002) (García-Borrón and Solano, 2002).

Supplemental Data S1: Pfam alignment results.

Shown below are the results of analyses of Pfam alignment of the Hidden Markov Models (HMM) of tyrosinase against E. dermatitidis tyrosinase sequences utilizing the Pfam tool. It should be noted from the outside that all *Exophiala dermatitidis* tyrosinase sequences match well to the tyrosinase HMM.

***Exophiala dermatitidis* tyrosinase sequence XP_009160170.1**

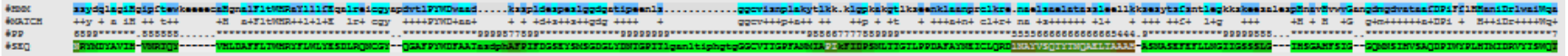
Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites
				Start	End	Start	End	From	To				
Tyrosinase	Common central domain of tyrosinase	Domain	CL0205	78	324	79	323	2	221	222	169.8	1.1e-49	n/a



Note that this sequence also provided a poor match to the Tyrosinase C HMM.

***Exophiala dermatitidis* tyrosinase sequence XP_009156893.1**

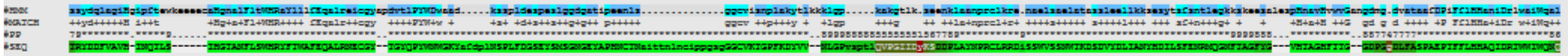
Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites
				Start	End	Start	End	From	To				
Tyrosinase	Common central domain of tyrosinase	Domain	CL0205	91	316	92	315	2	221	222	144.1	7.7e-42	n/a



Note that this sequence also provided a poor match to the Tyrosinase C HMM.

***Exophiala dermatitidis* tyrosinase sequence XP_009157733.1**

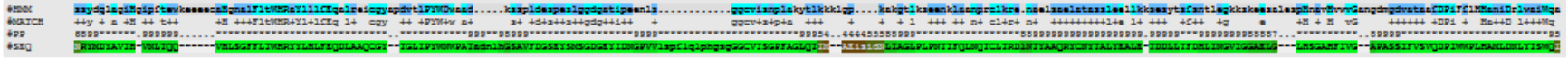
Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites
				Start	End	Start	End	From	To				
Tyrosinase	Common central domain of tyrosinase	Domain	CL0205	116	348	117	348	2	222	222	184.6	3.1e-54	n/a



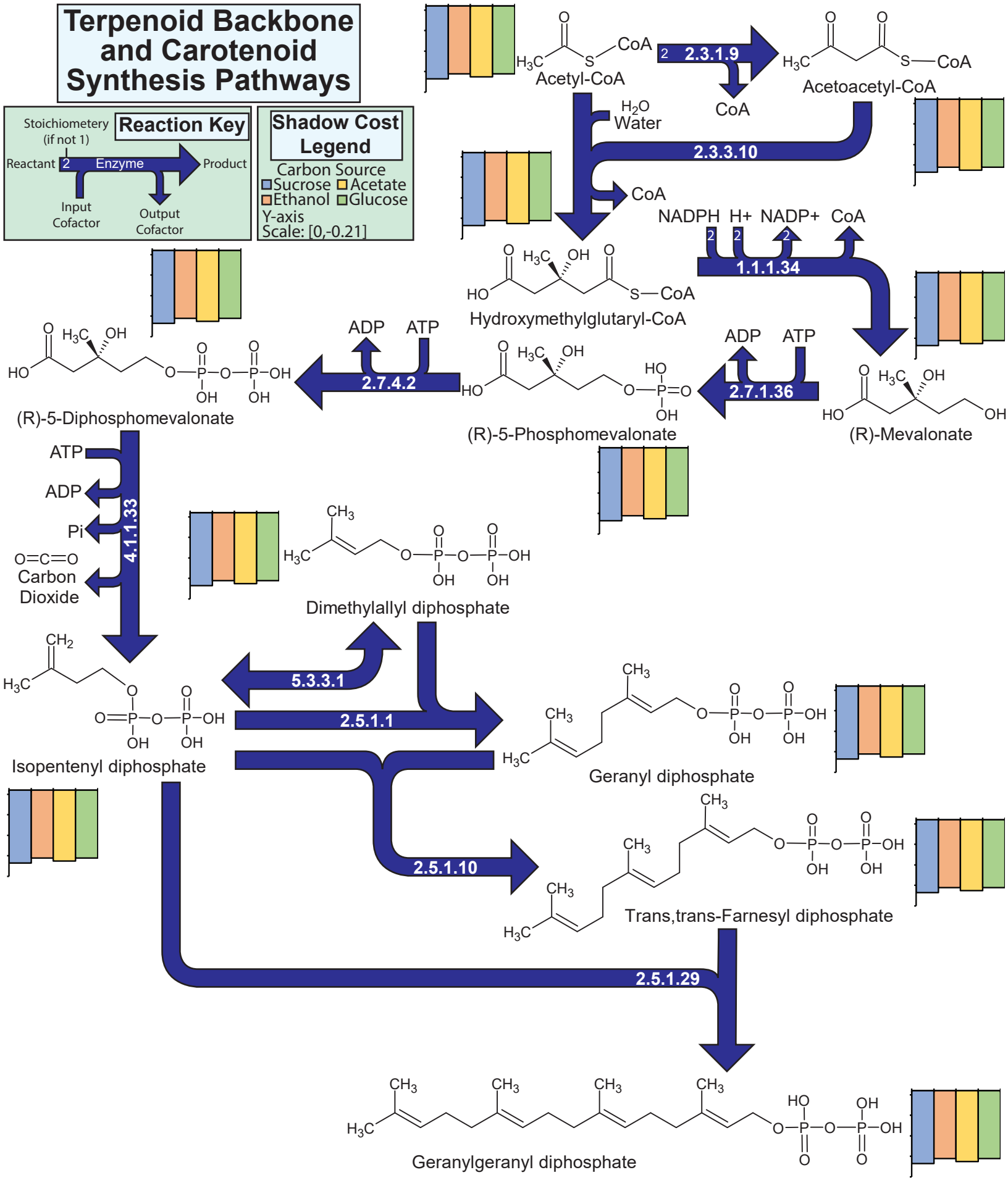
Note that this sequence matched to no other protein family HMM.

***Exophiala dermatitidis* tyrosinase sequence XP_009155657.1**

Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites
				Start	End	Start	End	From	To				
Tyrosinase	Common central domain of tyrosinase	Domain	CL0205	94	320	95	319	2	221	222	132.9	2.0e-38	n/a

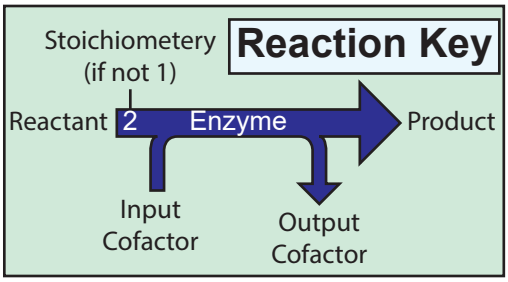


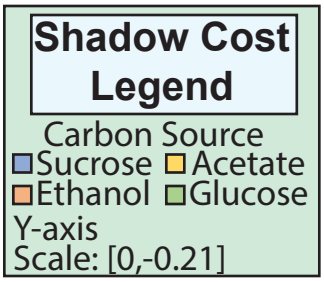
Note that this sequence matched to no other protein family HMM. Also note that this is the poorest match of any *E. dermatitidis* tyrosinase sequence to the HMM tyrosinase sequence, even through the e-value of the match is well below that which has been used as a cutoff throughout this work (1E-30).

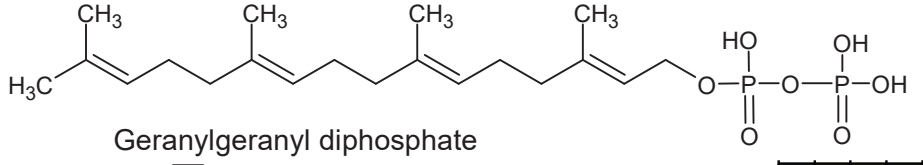


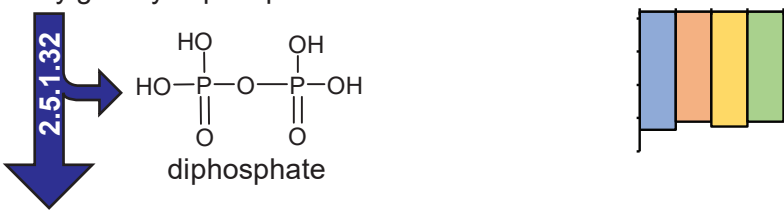
Supplemental Figure S1

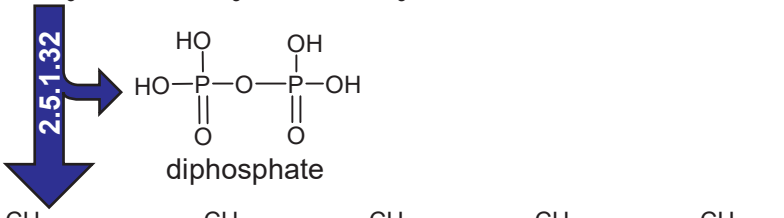
Carotenoid Synthesis Pathways

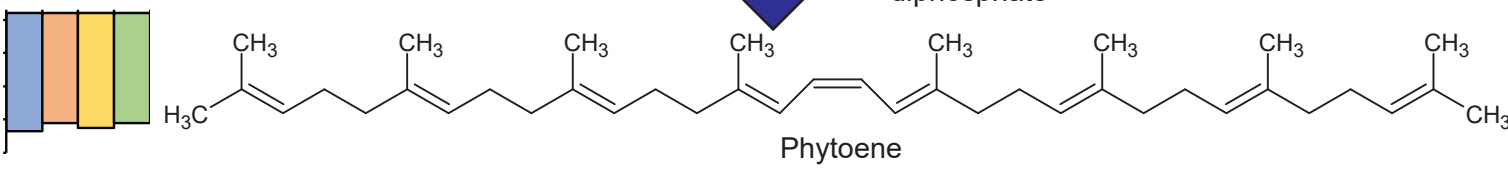


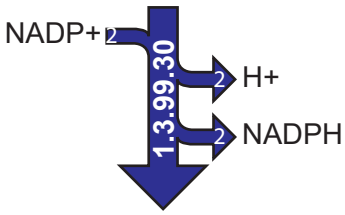


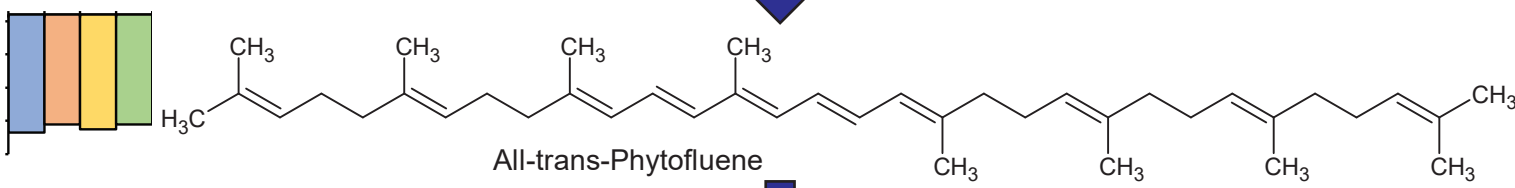


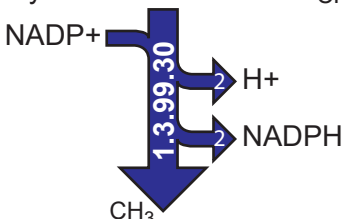


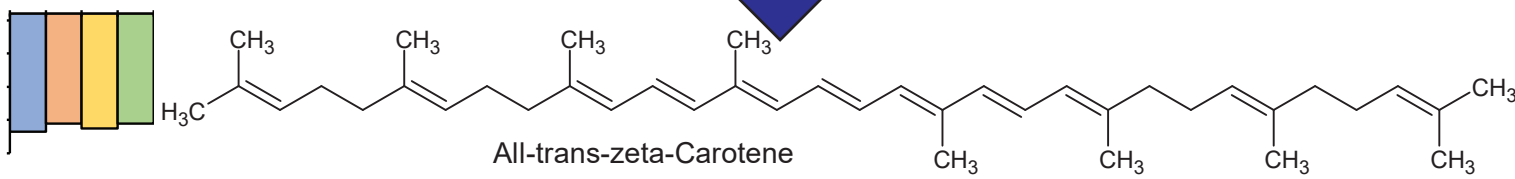


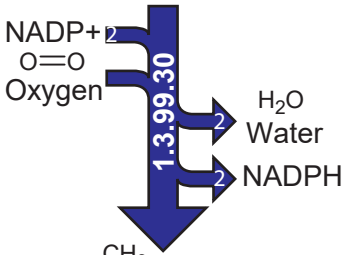


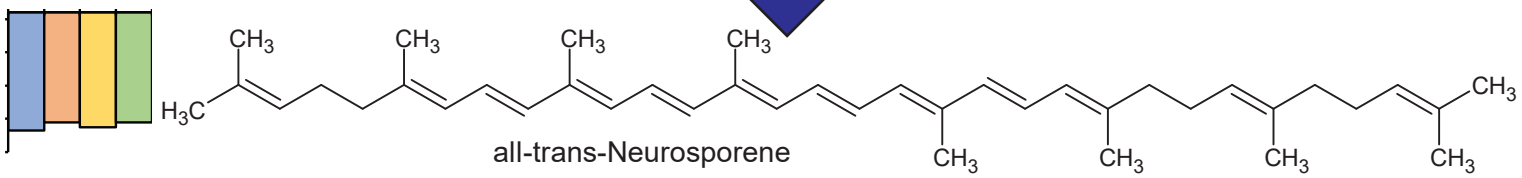






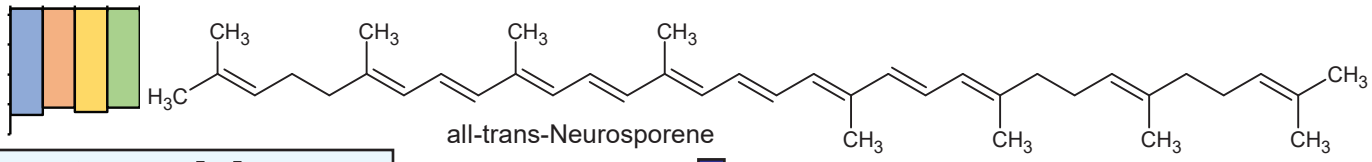


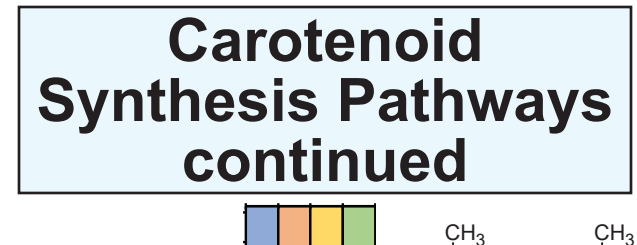


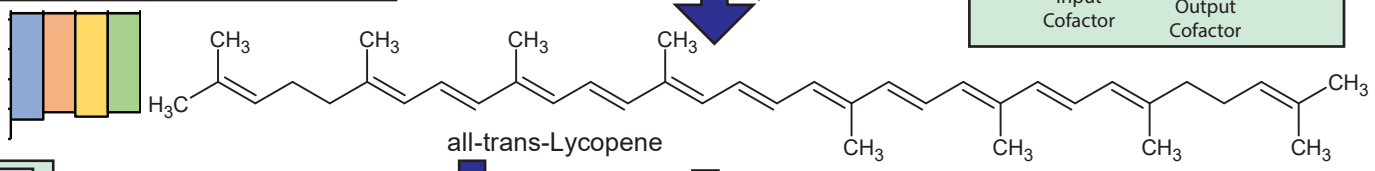


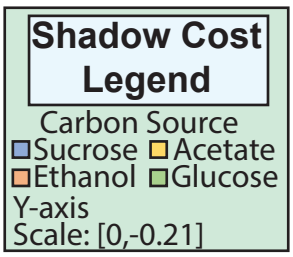
Supplemental Figure S2

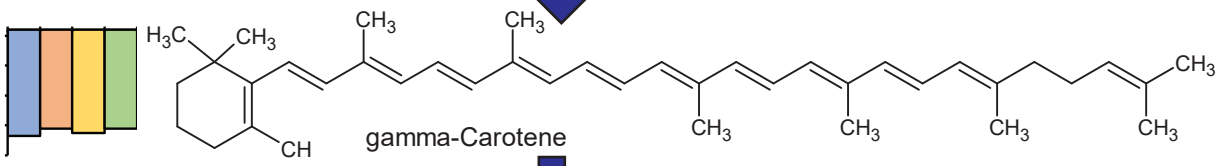
Carotenoid Synthesis Pathways continued

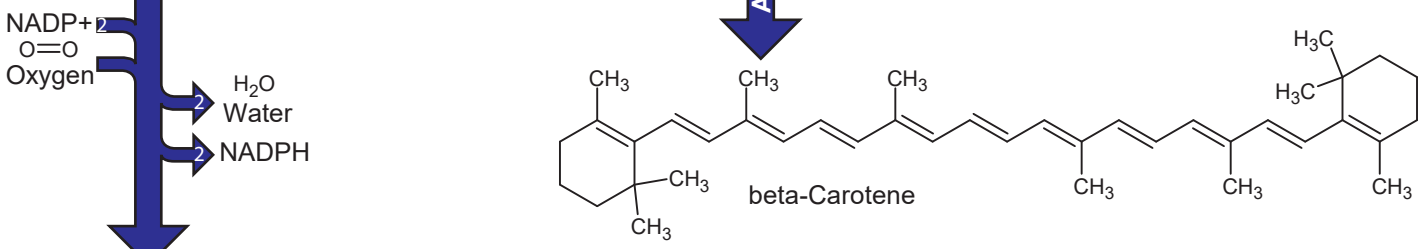


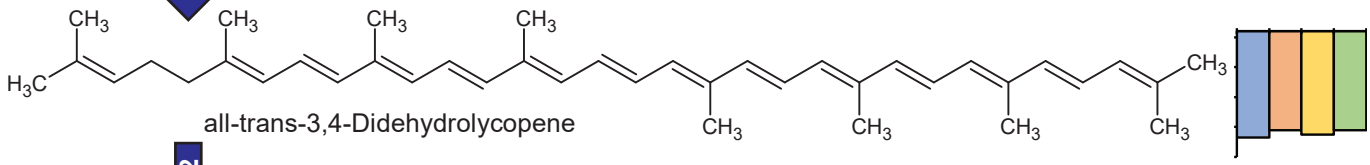


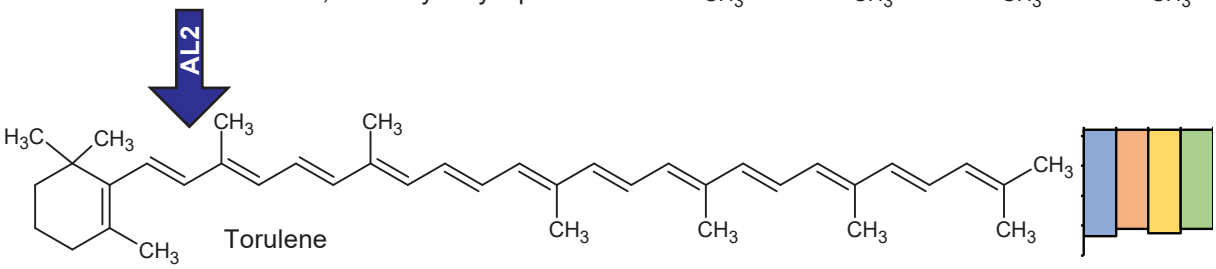


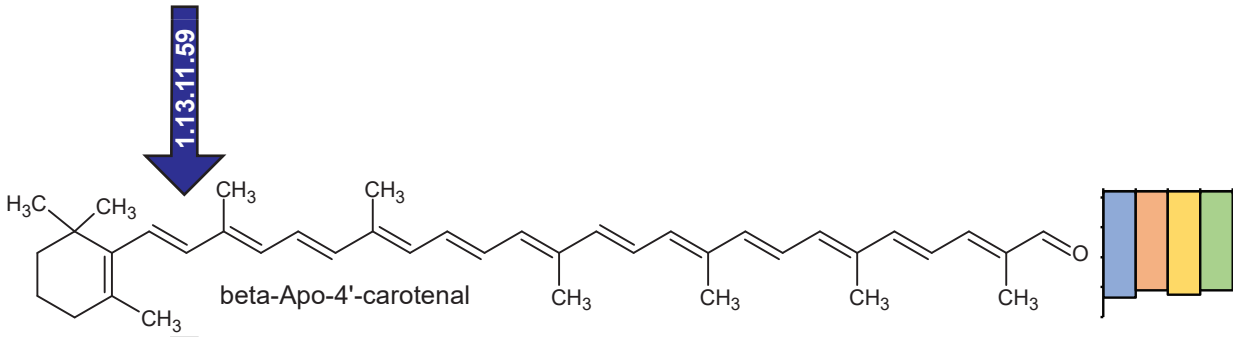


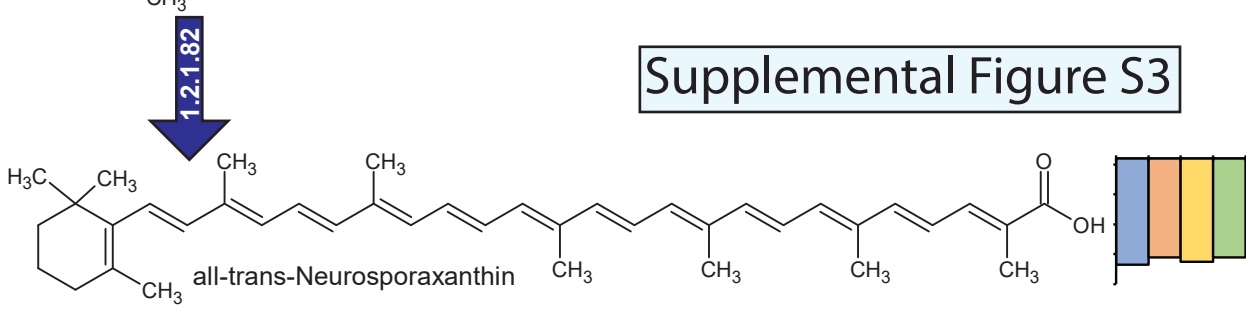












Supplemental Figure S3